A COMPARATIVE STUDY OF BASIC COLLEGE GRADES AND EFFORT-INTEREST-ATTITUDE RATINGS FOR LOW ABILITY STUDENTS

By

Leonard S. Laws

A THESIS

Submitted to the School of Graduate Studies of Michigan State College of Agriculture and Applied Science in partial fulfillment of the requirements for the degree of

DOCTOR OF EDUCATION

Department of Higher Education

1953

A COMPARATIVE STUDY OF BASIC COLLEGE GRADES AND EFFORT-INTEREST-ATTITUDE RATINGS FOR LOW ABILITY STUDENTS

By

Leonard S. Lewis

ABSTRACT

Submitted to the School of Graduate Studies of Michigan State College of Agriculture and Applied Science in partial fulfillment of the requirements for the degree of

DOCTOR OF EDUCATION

Department of Higher Education

1953

Approved Milosh Muntyan

This study was primarily concerned with the establishing of pass-fail points in the grading of Basic College students on a basis of the effort being made by low-ability students. The general hypothesis proposed by the Basic College Educational Research Committee was that "among a group of students, all low in ability, effort should be a significant factor differentiating those who pass from those who fail."

The low ability group was composed of 120 freshmen who received low scores on the 1949 edition of the American Council on Education Psychological Examination during the Fall Quarter, 1952 orientation program and who failed to increase their score appreciably on a re-test. Information relative to the students' studying was obtained by means of questionnaires sent to the students themselves, to dormitory assistants living in the same housing units as the students, and to instructors having the students in their classes. Each instructor rated his students twice during the quarter.

The study was restricted to first-quarter subjects being offered by the Basic College during fall term. These were Basic 111, first in a year's sequence in Communication Skills, and Basic 121, first in the Natural Science sequence.

Pooled ratings of the students' effort were obtained and correlations between the pooled ratings and grades received in Basic 111 and Basic 121 were computed. There was such a weak relationship between the grades and pooled effort ratings that it was not possible to establish the desired pass-fail points. Comparisons of the effort ratings made by the various observers suggested that they have little in common and might be too crude for discriminating between passing and failing students. More refined instruments are recommended for future studies.

There seemed to be little relationship between the prediction by the American Council on Education Psychological Examination and the grades the students received. Many capable students were included in the low ability group - students who were not in danger of failing regardless of effort. Better prediction will be necessary if future studies of this type are to succeed.

The Basic College has a general policy that: "Instructors' grades and examination grades should meaningfully supplement, rather than duplicate each other." A second purpose of this study was to investigate whether such meaningful supplementation is being evidenced. This investigation utilized ratings on the students' general

attitude and ratings on the students' ability that were made by the Basic 111 and Basic 121 instructors. It further utilized the students' stated interest in the respective courses.

The evidence indicates that there might be meaningful supplementation with regard to factors involved in the students' general attitude. However, in the case of instructors' judgments as to students' ability to meet course requirements there appears to be supplementation but not meaningful supplementation. It is recommended that a more extensive study be undertaken to discover wherein the instructors' grades might more meaningfully supplement final examination grades.

- ACKNOWLEDGMENTS -

The author wishes to express his sincere thanks to Dr. Walker H. Hill for having consistently and generously given himself completely to the careful consideration of problems which arose in connection with this study. His suggestions have been an invaluable aid in the development and writing of this dissertation.

He is also greatly indebted to Dr. Milosh Muntyan whose cooperation and encouragement have made this study possible, and whose constructive criticisms have greatly improved the quality of this thesis.

Special mention is deserved by Dr. Cecil V. Millard, Dr. Vernon G. Grove, and Dr. Harry Sundwall for their kind guidance and helpful suggestions.

Grateful acknowledgment is made for the cooperation of Dean C. E. Erickson, Dr. Paul L. Dressel, Prof. Paul D. Bagwell, and Dr. Chester A. Lawson. Their approval and support greatly facilitated the gathering of data necessary for this study.

The author is grateful to Dr. Willard G. Warrington for his helpful suggestions relative to statistical techniques and to Dr. Robert A. Jackson for pertinent data that he generously supplied.

Sincere thanks are extended to the many individuals who supplied the data - the instructors in the Departments of Communication Skills and Natural Science, the dormitory assistants, and the students who participated in this study. Their friendly cooperation was vitally necessary and greatly appreciated.

The author would have been unable to finance his doctoral program, including this study, had he not received the graduate assistantships and tuition scholarships that have been so generously bestowed by Michigan State College. For these grants-in-aid he will be forever grateful.

This dissertation is dedicated to Janet Owen Laws whose patience, confidence, and devotion have fostered the continuous concentration necessary to the successful completion of this study.

- VITA -

Leonard Stewart Laws

candidate for the degree of

Doctor of Education

Final Examination, May 20, 1953, 2:00 p.m. Room 201, Morrill Hall

Dissertation: A Comparative Study of Basic College Grades and Effort-Interest-Attitude Ratings for Low Ability Students

Outline of Studies

Major Subject: Higher Education

Cognate Area: Mathematics

Biographical Items

Born, December 29, 1917, Pocasset, Oklahoma

Undergraduate Studies, Willamette University, 1935-1939

Graduate Studies, Leland Stanford Jr. University, 1939-1941
University of Minnesota, 1941-1942
Cont. 1945-1946
Michigan State College, 1947-1948
Cont. 1952-1953

Experience: Teaching Assistant, Stanford University, 1939-1941, Teaching Assistant, University of Minnesota, 1941-1942, Instructor, University of Minnesota, 1942-1945, Assistant Professor, University of Minnesota, 1945-1947, Graduate Assistant, Michigan State College, 1947-1948, Assistant Professor, University of Minnesota, 1948-1952, Graduate Assistant, Michigan State College, 1952-1953.

Member of Kappa Delta Pi

TABLE OF CONTENTS

TABLE OF TABLES

TABLE OF TABLES
(continued)

TABLE OF HISTOGRAMS AND GRAPH

CHAPTER I

INTRODUCTION

Michigan State College has traditionally emphasized technical training. Established in 1855 as the first of the state colleges of agriculture, it became, under provisions of the Morrill Act, one of the nation's land grant colleges. It has tried through the years to meet the needs of its people and to provide them with a useful education.

The original Agriculture College of Michigan included "Political Economy" and "Natural Philosophy" in its curriculum along with surveying and leveling and the more strictly agricultural subjects. It stressed the use, understanding, and appreciation of the English language.[1]

In recent years, under the inspiring leadership of President John A. Hannah, Michigan State College has recognized the need for giving all undergraduate students a greater breadth of educational experience than that attained in the prerequisite courses of the various technical fields. As a result of this recognition, and consistent with its established tradition of providing a useful education to meet the needs of its people, Michigan State

---

[1] Dressel, Paul L. and Others. Comprehensive Examinations in a Program of General Education. Michigan State College Press, East Lansing. 1949. 165 pp.

College developed a program of general education having the following broad objectives:

"The following outcomes are considered not only useful, but also fundamental to a sound educational experience for all Michigan State College students: to be able to write and speak clearly, concisely, and effectively, and to be able to understand, appreciate, and evaluate critically the written and spoken word; to know something of the biological and physical sciences, not merely from the professional's viewpoint, but more particularly in respect to their impacts on everyday living and thinking; to have an interest in, and a knowledge of, personal, family, social, and civic affairs; to be acquainted with the facts of history, particularly the history of selected periods most significant in relation to the world of today; to have an appreciation of the cultures, past and present, expressed in literature, music, and art."[2]

The Basic College, the administrative unit responsible for the general education program of Michigan State College, was organized in 1944 to help all students of Michigan State College attain the objectives stated above. As originally organized, it included seven departments which were reorganized into four departments beginning with fall quarter,

---

[2]Ibid, p.1

1952, the period of this study. The four organizational areas are now: Communication Skills, Natural Science, Social Science, and Humanities. This study will be concerned with two of these, namely: Communication Skills and Natural Science.

Michigan State College recognizes the need in present-day society for highly trained specialists, but at the same time feels that there is a core of educational experience that should be shared by all students, regardless of their special interests.

Howard C. Rather, former Dean of the Basic College has stated: "The Basic College Program is designed to make sure that no undergraduate at Michigan State College follows a specialized program so intensively that in the end he knows too little of the relation of his own specialty to other activities and to the needs of society as a whole. It is designed to build specialized training, where desirable, on a broader foundation. It is designed to give each student - whether he be an eventual specialist or not - the opportunity for knowledge, skill, understanding, appreciation, and thinking in diverse ways, so that he may develop as a well rounded individual, capable of adjustment to changing conditions; capable not only of

rendering service on a job, but also of utilizing effectively those nonwork, nonsleep hours that constitute so important a part of the good life."[3]

Dean Rather also pointed out that the students starting this program of general education are being taught by staff members who are trained and primarily interested in general education under an administrative unit that recognizes and promotes good teaching with rewards in salary, rank, and professional advancement fully equal to those granted for research, writing, or any other educational activity.

While the total enrollment in the Basic College has been large, special effort has been made to keep the classes small to provide each instructor an opportunity for becoming better acquainted with his students. The instructors whose ratings are reported in this study had thirty students or less in their classes.

As the Basic College was originally organized, the students' grades in their Basic courses were determined solely on a basis of their performance on comprehensive examinations. These examinations covered the full year's work (Fall, Winter and Spring quarters) for each course of study. The comprehensive examinations were developed

---

[3]Ibid, p.3

cooperatively by a committee from the department concerned and an independent Board of Examiners.

While such an examining and grading system has much to be said for it, there has continued to be some feeling on the part of the faculty that quarter by quarter grades should be given and that the grade should be partially determined by the instructor's appraisal of the student's ability and performance. Such a system has recently been adopted for the Basic College. During fall quarter, 1952, the period of this study, the instructor's grade had equal weighting with the final examination grade in the determination of the course grade.

Under its general policy for the determination of grades, the Basic College has listed:

> "The instructor's grade will be based on such evidence as is, in his judgment, appropriate and is in accordance with policies determined by his department and/or the dean.
>
> Instructors' grades and examination marks should meaningfully supplement rather than duplicate each other.

and

> The distribution of term grades should conform as closely as possible to the percentage distribution:
>
> A 7-11 B 25-29 C 45-49 D 12-16 F 0-5."

Because the percentage distribution of scores was arbitrarily determined and because studies of the actual distributions of scores indicated that grading practices in the Basic College might not be as generous as those in comparable educational units elsewhere, Dr. C. E. Erickson, present Dean, requested that the Basic College Educational Research Committee "give some study to the problem of assignment of marks with the purpose in mind of arriving at a plan which might replace the interim one of assigning marks according to a specified distribution."[4]

The Research Committee first established the most important functions or purposes of marks, the major desirable qualities of a mark, and the major bases for determining marks as seen by the faculty of the Basic College. It then undertook three essentially different studies to accumulate evidence upon the problem of assign-

---

[4]Basic College Educational Research Committee. The Assignment of Term Marks in the Basic College. Mimeographed Report. Michigan State College. 13 numb. leaves.

ing term marks.

Each study attempted to answer one of the following questions:

1. "Are students in Basic courses working anywhere close to the level which might be expected of them in terms of their ability?
2. How does the distribution of marks given in Basic courses at M.S.C. compare with distributions for other freshman and sophomore courses, both at M.S.C. and at other similar institutions?
3. Is there agreement among Basic College staff members of a given Basic course as to the level of achievement which should characterize a given letter grade?"[5]

In trying to answer the first question, three approaches were made. One was by referring to residence hall reports to locate students reported as not working. "The assumption on which this approach was based was that students of low ability reported as not working should usually receive D's or F's in Basic courses. If not, and

---

[5]Basic College Educational Research Committee. The Assignment of Term Marks in the Basic College. Mimeographed Report. Michigan State College. March 195[illegible]. 9 numb. leaves.

if the residence hall reports on this point have much validity, it would seem that the average level of achievement is too low. One study was made on this point for students of Spring (Sic) 1951, but the rating on effort used (Sic) was rather informal and unreliable. Hence, a request was made to the residence hall staff for cooperation on a special report, twice during the Winter Term 1952, of those individuals deemed to be making no scholastic effort. These reports were obtained, analyzed, and the results separately reported in a mimeographed statement. Essentially it was found that there was no meaningful relationship between residence hall reports on work or lack of it by students and the grades obtained in Basic courses.

"A more exacting study was undertaken during Fall 1952. This approach involved the obtaining of reports from residence halls, from instructors, from the Counseling Center, from Improvement Services, and from the students themselves as to the effort made. The general hypothesis was that among a group of students, all low in ability, effort should be a significant factor differentiating those who pass from those who fail."[6]

---

[6]Ibid.

The study with which this thesis is concerned undertook to investigate for the Basic College Educational Research Committee whether a pass-fail point might be established for Basic College grading purposes on a basis of the effort made by students to meet their course requirements, as reported by all available campus sources of information.

A further purpose of this dissertation is to show what relationships exist between the marks given in the Basic 111 and Basic 121 courses and factors such as the student's stated interest in the course, his apparent ability to meet the course requirements, his general attitude toward the course as appraised by his instructor, and the effort that he put forth to master the material of the course as judged by various observers.

The Basic College Educational Research Committee found the major factors listed by Basic College teachers as entering into the determination of their grades to be:

1. "Knowledge of the specific materials covered in the course.
2. Proficiency, intellectual skills, or abilities listed as course objectives.
3. The improvement made by a student during the course.

4. Attitudes toward society, the course, or the instructor demonstrated by behavior, effort, enthusiasm, and cooperation on the positive side, or by inattention, lack of cooperation, or absence on the negative side."[7]

The Research Committee felt that the first two factors were more likely to possess the desirable qualities of "objectivity, reliability, uniformity, and validity"[8] than other factors in the complex that enters into decisions on marks.

This dissertation is primarily interested in a factor, namely the students' study effort, that is not included in the first two considerations listed above. It has as a secondary purpose the study of other such factors that enter into the determination of the instructor's mark.

One of the specifications for determining grades in Basic College was: "Instructors' grades and examination

---

[7]Basic College Educational Research Committee. The Assignment of Term Marks in the Basic College. Mimeographed Report. Michigan State College. 13 numb. leaves.

[8]Basic College Educational Research Committee. The Assignment of Term Marks in the Basic College. Mimeographed Report. Michigan State College. March 1953. 9 numb. leaves.

grades should meaningfully supplement rather than duplicate each other."[9] This study investigates whether personal factors such as the student's stated interest in the course, the instructor's appraisal of the student's general attitude toward the course, and the instructor's opinion as to the student's ability to meet the course requirements enter into the instructor's mark to a greater extent than they do to the final examination score. To do this, each factor must be studied in turn and many relationships and inter-relationships between factors must be examined to discover the composition of the instructor's mark if that can be found.

Answers to the following list of questions will aid in determining whether instructors' marks "meaningfully supplement" the final examination grades.

1. Is the instructor's appraisal of the student's general attitude toward the course related to:
   (a) the final examination score?
   (b) the instructor's end-of-quarter mark?
   (c) the student's stated interest in the course?

---

[9]Basic College. Policies and Procedures for Term End Examinations and Term Grades in the Basic College. Unpublished Booklet. Michigan State College. 13 numb. leaves.

2. Is the instructor's estimate of the student's ability to meet the course requirements related to:

    (a) A.C.E. Test-retest Average of the student?

    (b) similar estimate by an instructor in another subject matter field?

    (c) a second estimate of the same thing by the same instructor?

    (d) the final examination score?

    (e) the instructor's end-of-quarter mark?

    (f) the ability level of the student?

    (g) the student's apparent attitude toward the course?

    (h) the mid-term grade given by the instructor?

3. Is the instructor's rating of the student's effort toward mastering the skills and materials of the course related to:

    (a) a second rating of the same factor by the same instructor about the same student?

    (b) the final examination score?

(c) the instructor's end-of-quarter mark?

(d) the number of hours per week the student states that he has spent studying the subject?

4. Is the student's stated interest in the course related to:

(a) the final examination score?

(b) the instructor's end-of-quarter mark?

(c) the ability level of the student?

5. Are there factors entering into the instructor's marks that are not meaningful in supplementing the final examination grades?

Communication Skills (Basic 111-112-113) and Natural Science (Basic 121-122-123) were chosen for the purpose of this study because they were the two Basic College courses normally taken by entering freshmen. Since the study was primarily interested in the low-ability students - students who might be barely passing or in danger of failing, and since some of these students might not be enrolled for more than their first quarter, it was deemed necessary to confine the study to the first quarter's work. Hence, the study was restricted to Basic 111 and Basic 121.

This proved to be a fortunate choice, since in those two areas there is greater opportunity for the instructors to become acquainted with the individual students than in either Social Science or Humanities courses where classes are considerably larger.

More than 3000 students were enrolled in Basic 111 during fall quarter, 1952. These students were registered for a one-hour lecture session each week, accommodating many students, and were also registered for four hours per week in small laboratory sections devoted to discussion, reading, writing, and speaking. There were more than one hundred laboratory sections and the enrollment in each was limited to thirty or fewer students. Some instructors had only one of these sections, while other instructors were in charge of two, three, or four sections each. Each instructor was assigned by his Head of Department to teach his particular sections and that assignment was made without knowledge as to which particular students might be in the sections concerned.

The students were signed into the sections without regard to which instructors might be in charge. Thus the student's interest in the course was not conditioned by his prior knowledge as to the instructor that he might have.

The arrangements in Basic 121 were similar to those in Basic 111 with the exception that the students spent two hours per week in the large lecture sections and three hours per week in laboratory.

While the over-all course might have been the same for the students enrolled in Basic 111 (or Basic 121), the particular experience of one student in one section under one of the teachers would differ in many respects from the educational experience of another student either in that same section or in another section and under another instructor. These environmental differences will undoubtedly influence the findings of this study.

CHAPTER II

DEFINITION AND SELECTION OF LOW-ABILITY STUDENTS

The problem of establishing a pass-fail point by comparing the grades received by low-ability students who were definitely studying to those received by low-ability students who were making little effort demanded that extra care be exercised in selecting the low-ability students. They should be students of such low academic ability that they would be in danger of failing if they made little effort to succeed.

The problem could have been attacked by first discovering which students failed and which barely passed, then getting information about the studying they did during the previous quarter. This approach was rejected because of the difficulty in getting reliable information in retrospect.

The method chosen for this study was to first predict as accurately as possible which students might be low in academic ability, then gather information about their study effort while they were making that effort. This information might then be compared to the grades the students received at the end of the quarter. The greatest obstacle to this approach is in obtaining a reliable and valid prediction of the student's ability to meet the academic requirements of the Basic College courses.

Thus far there is no perfect predictive instrument available for determining academic success. The Basic College

Educational Research Committee in collaboration with the Board of Examiners of Michigan State College decided to use the 1949 edition of the American Council on Education Psychological Examination as the predictive instrument for this study. This decision was influenced by the availability of scores on that particular examination since all students are required to take it upon entering Michigan State College as part of the regular admissions procedure.

Although the A.C.E. Psychological Examination is widely used for indicating college aptitude, it was decided to be cautious in selecting the low-ability group by using it twice instead of the usual once. Only the students that remained low on the second administration of the test would be included in the sample group and called low-ability. While it was recognized that this would not guarantee the selection of all lowest-ability students (by academic standards), it was the best that could be done at the time with the instruments available.

The Research Committee realized that by selecting low-ability students in this manner it might not find a sharp cleavage of grade between those students who were trying to succeed and those who were not. However, it felt that as a group the students making a sincere effort should get better grades than the low-effort students.

The range of A.C.E. scores on all fall quarter, 1952 freshmen was divided into ten equal units. Each unit was approximately 0.6 of one standard deviation, under the assumption of a normal distribution being dispersed ± 3 sigma about its mean. As the units were called standard scores, the entire distribution was included in a range of ten standard scores. This method of reporting scores in terms of standard scores was common procedure for the Board of Examiners at Michigan State College.

Those freshmen receiving A.C.E. Psychological scores in the lowest four standard scores of the entire distribution of scores were requested to retake the examination at a later date. This group, approximately 28 percent of the entire freshman class, comprised those students having a raw score of 82 or below on the 1949 edition of the A.C.E. Psychological Examination. If their raw scores were 85 or less on the second administration of the same test, they were considered to be low-ability students for the purpose of this study. The low-ability group also included thirty students who did not take the Psychological test the second time, but whose first scores were so low that it was very unlikely for them to raise the A.C.E. Psychological score appreciably on a second administration

One hundred and thirty-nine students were selected by the test-retest method described above, but nineteen were eliminated for one of the following reasons:

1. It was desired to compare ratings and grades in Basic lll with corresponding ratings and grades in Basic 121. Those students not enrolled for both subjects were dropped from the study.
2. It was desired that appraisals of the student's effort be given by the dormitory assistants. Students not living in college dormitories were also excluded from the study.
3. Those students who dropped from college during the first quarter were also dropped from the study since it was impossible to obtain their grades and to obtain certain ratings.

While this latter group might have included lowest-ability students or students who quit when they encountered some difficulty, there were so few (3) of them that excluding them would not noticeably affect the results of the study. In fact, there is no reason to believe that results based upon the original group of 139 students would differ significantly from results based upon the 120 students remaining in the study.

Any results of this study are to be interpreted in light of the group composition. "Low-ability group" or "low-ability

students" as used in this study refer to students who received A.C.E. Psychological Examination scores that were in the lowest four standard scores of the total freshman distribution and who did not improve that score sufficiently on a second administration of the same test to be excluded from the group.

It should be pointed out that these low-ability students did receive grades in Basic 111 and in Basic 121 that statistically were significantly (5% level) lower than the grades received by the entire student population taking those courses. However, the low-ability group included many students who received C or better grades in the courses concerned. In fact, 51.3 percent of the low-ability group received C or better grades in Basic 111 and 41.1 percent did the same for Basic 121. This emphasizes the fact that while the group as a whole did more poorly than freshman students generally, many in the group would be considered good students. Because of this and other considerations to be discussed later in this dissertation, conclusions based upon the evidence to be presented might apply to the freshmen students generally rather than to the low - ability group exclusively.

These findings relative to the effectiveness of the American Council on Education Psychological Examination as an instrument for predicting first quarter scores or grades are in general agreement with studies made in other colleges. Some of these findings are as follows:

Wagner and Strabel[1] report a correlation of 0.22 between A.C.E. Psychological test scores and the freshman-sophomore physical science average on 651 students enrolled at the University of Buffalo during the 1925-1929 period. They also report a correlation of 0.40 between freshman English grades and A.C.E. scores for 661 University of Buffalo students.[2]

Quaid[3] reports a correlation of 0.408 between the first semester average of college freshman marks at Philips University (1934-1935) and the A.C.E. scores for 140 students.

Leaf[4] showed the A.C.E. to be the poorest predictor of college success of all instruments used in his study.

Froehlich[5] reports a correlation of 0.554 between A.C.E. scores and grade point averages on the first quarter's work for 1316 regular students at the University of Wisconsin.

---

[1]Wagner, Mazie E. and Strabel, E. Predicting Success in College Physical Sciences. Science Education 19:4-9. 1935.

[2]Wagner, Mazie E. and Strabel, E. Predicting Performance in College English. Journal of Educational Research 30:694-9. 1937.

[3]Quaid, T. D. D. A Study in the Prediction of College Freshman Marks. Journal of Experimental Education 6:350-75. 1938.

[4]Leaf, Curtis T. Prediction of College Marks. Journal of Experimental Education 8:303-7. 1940.

[5]Froehlich, Gustav J. Academic Prediction at the University of Wisconsin. Journal of American Association of College Registrars 17:65-76. October 1941.

Brown[6] states "--The students below the 20th percentile on total (A.C.E.) score would not ordinarily be considered good college prospects, but almost half of them achieve a grade point average of C or better." He reports correlations between the A.C.E. scores and grade point averages in "Quantitative Subjects" such as mathematics and science to be 0.40. Between the A.C.E. scores and grade point averages in "Linguistic Subjects" such as English, social sciences, and languages, he found a correlation of 0.44. His correlation between total A.C.E. scores and total grade point averages on all subjects turned out to be 0.40 for the 124 members of his random sample selected from 1048 entering freshmen in the Liberal Arts Division of Long Beach City (Junior) College.[7]

Osborne, Sanders, and Greene[8] found that "success in certain subject matter areas is predicted with markedly greater accuracy than in other subjects. The higher r's are found for the natural sciences and languages."

A further study reported by Jackson[9] indicates a correlation of 0.47 between the A.C.E. and first-quarter grade point

---

[6]Brown, Hugh S. Differential Prediction by the A.C.E. Journal of Educational Research 44:116-121. October 1950.

[7]Brown, Hugh S. Differential Prediction by the A.C.E. Journal of Educational Research 44:116-121. October 1950.

[8]Osborne, R. Travis, Sanders, W. B. and Greene, J. E. The Differential Prediction of College Marks by A.C.E. Scores. Journal of Educational Research 44:107-115. 1950.

[9]Jackson, Robert A., Member of Michigan State College Board of Examiners (Oral Communication).

averages for 2832 Michigan State College students based upon fall quarter grades, 1952. Jackson's group included many, if not all, of the students in this study.

These studies suggest that by using the American Council on Education Psychological Examination as a criterion of low ability it might not be possible to get a group of students who are truly low in academic ability - students who will be in danger of failing unless they put forth great effort. It is evident that members of the sample selected for this study were not all in danger of failing. This fact will undoubtedly lessen the likelihood of establishing a pass-fail point on a basis of student effort.

While the problem of predicting which students were truly low in academic ability was of vital importance to the part of this study concerned with the establishment of a pass-fail point, it was not crucial to other parts of the study. For instance: whether a student has low ability or not and whether the instructor can accurately estimate the student's true ability or not, it is still of interest to discover the extent to which the instructor's rating of the student's ability is reflected in the mark the instructor gives that student for his quarter's work.

CHAPTER III

MEASURABILITY OF THE VARIABLES

Three factors must be considered if the data of this study are to be properly interpreted. These factors are:

1. The nature of the variables to be measured.
2. The measuring instruments and methods of application.
3. The methods used for analyzing the data with the assumptions and qualifications warranting their use.

This chapter is primarily concerned with the first factor in the list above. The remaining factors will be discussed in successive chapters.

Thorndike has stated "Anything that exists at all exists in some amount. To know it thoroughly involves knowing its quantity as well as its quality. Education is concerned with changes in human beings; a change is a difference between two conditions; each of these conditions is known to us only by the products produced by it--things made, words spoken, acts performed, and the like. To measure any of these products means to define its amount in some way so that competent persons will know how large it is, better than they would without measurement. To measure a product well means so to define its amount that competent persons will know how large it is, with

some precision, and that this knowledge may be conveniently recorded and used."[1]

Variables such as a student's interest in a course of study, his general attitude toward that course, the effort that he makes to succeed in the course, and his ability to succeed are all recognized by teachers as existing. Not only do teachers recognize the existence of these variables as an observable fact but see them as existing in differing amounts in different students or in the same student at different times.

The fact that instructors in Basic 111 or Basic 121 agree as to the existence of a particular variable such as a student's attitude toward their course does not imply that two or more instructors in the same subject matter area would agree as to the quality of a particular student's attitude - even though they were all to observe him in the same classroom and at the same time. Each instructor might have his own standard for rating the student on the variable under observation.

Concepts such as a student's general attitude toward a course are actually multi-dimensional. A student might be rated as having a good attitude because he appears interested

---

[1]Thorndike, E. L. The Nature, Purpose and General Methods of Educational Products. 17th Yearbook, National Society for the Study of Education. Part 2. p. 16.

in the class discussions and in the recitations. Or, he might get the same rating because he asked thoughtful questions - seemed to have insight into the problems of the course. Or, again, he might seem eager to learn, or pay rapt attention to what the instructor says. All of these and many others are physical manifestations of the student's attitude toward the course as observed by the instructor concerned. And yet in spite of this multi-dimensionality, it is common practice for instructors to regard a student as having a good (or other) attitude toward his course. Apparently, students are compared in the minds of their instructors according to a linear scale. This leads to the question: How well can this multi-dimensional variable be measured with a one-dimensional or linear rating scale?

Thurstone and Chave[2] had a similar problem in measuring attitudes. They argued as follows: "When the idea of measurement is applied to scholastic achievement, for example, it is necessary to force the qualitative variations into a scholastic linear scale of some kind. We judge in a similar way qualities such as mechanical skill, the excellence of handwriting, and the amount of a man's education as though these traits were strung out along a single scale, although they are, of course, in reality scattered in many dimensions. As a matter of fact, we get along quite well with the concept

---

[2]Thurston, L. L. and Chave, E. J. The Measurement of Attitude. University of Chicago Press, Chicago, Ill. 1929. p. 10.

of a linear scale in describing traits even so qualitative as education, social and economic status, or beauty. A scale or linear continuum is implied when we say that a man has more education than another, or that a woman is more beautiful than another, even though, if pressed, we admit that perhaps the pair involved in each of the comparisons have little in common. It is clear that the linear continuum which is implied in a 'more and less' judgment may be conceptual, that it does not necessarily have the physical existence of a yardstick."

Apparently the instructor's linear rating of the student's class attitude is not requiring any more "forcing" than the instructor's linear rating of the student's total performance during the quarter as indicated by the grades A, B, C, D, or F. At most, the error in appraising general class attitude is of no greater magnitude than the error in appraising total performance. Even though more accurate instruments for measuring general class attitude might be devised than were used in this study, no increase in accuracy might be expected in the comparisons of ratings with grades unless the grading system itself was refined.

It should be noted that the ratings obtained are but indices of the true situation. The fact that an instructor felt that a student had a good class attitude does not imply that the student himself or that other instructors would feel the same way about it. If, for instance, an instructor

rates a student as having a good class attitude because the student appears to be eager to learn and is responsive to the ideas of the instructor, there is no guarantee that the student is actually that way.

"But this discrepancy between the index and 'truth' is universal. When you want to know the temperature of your room, you look at the thermometer and use its reading as an index of temperature just as though there were no error in the index and just as though there were a single temperature reading which is the 'correct' one for the room."[3]

Rating scales such as those used for gathering information for this study imply that the variable being observed is considered to be continuous - at least continuous in a certain range. For instance, the rating scale from the "Special Information Blank" (see Appendix):

Student's ability to meet the requirements of this course:

Very Low ——— Average ——— Very High

This scale implies that Ability might be rated as being somewhere from very low to very high, inclusive. Thurstone and Chave point out that such an assumption is rather common for measurement generally: "In almost every situation involving measurement there is postulated an abstract continuum such

---

[3] Ibid., p. 8.

as volume or temperature, and the allocation of the thing measured to that continuum is accomplished usually by indirect means through one or more indices. Truth is inferred only from the relative consistency of the several indices, since it is never directly known."[4]

The other variables of this study are also multi-dimensional and are assumed to be continuous. The statements above apply to them as well as to the instructor's appraisal of the student's ability.

In attempting to get at the true situation relative to the amount of effort being made by the students, this study has included as many indices as possible. Ratings were obtained from the Basic 111 and Basic 121 instructors having the low-ability students in their classes. The students themselves were asked to describe the studying that they did for their Basic College courses. Dormitory assistants were requested to complete a check list concerning the studying being done by those student members of the sample group that they were able to observe.

The dormitory assistants are undergraduate or graduate students who live in the dormitories and have the special responsibility of maintaining order, giving incidental help with studies, and being generally aware of all that is happening in the dormitory precinct (living unit) to which they are assigned by the residence counselor. The residence

[4] Ibid., p. 8.

counselors are faculty members living in the dormitories who have general supervisory and counseling duties.

By obtaining two ratings, at different times during the quarter, from the Basic 111 and Basic 121 instructors, it was possible to get two indices relative to the variables of apparent ability, general attitude, and apparent effort. It was thus possible to estimate the reliability of these types of ratings.

This study will attempt to discover the extent to which the various sources of information agree in their ratings of the student's effort. It will attempt to answer the question: Can the various indices be combined to give a meaningful index of the student's effort? Such a question must be investigated if a pass-fail point is to be established on a basis of pooled effort ratings.

CHAPTER IV

DESCRIPTION OF INSTRUMENTS AND PROCEDURES

Desired information was obtained by means of three questionnaires, copies of which are included in the Appendix. The "Special Information Blank" was submitted to the Communication Skills (Basic 111) instructors and to the Natural Science (Basic 121) instructors through their departmental offices at mid-term and again during the final week of classes in fall quarter, 1952.

The blanks were filled out as to Instructor's name, department identification, course and section numbers, and the student's name was written in the proper space before they were submitted to the instructors concerned. They were separated by instructor, by course, and by section; and alphabetically arranged within each section so as to require a minimum of handling by the staff members. This was done to impose as little as possible upon the time of the staff and thus gain greater cooperation with accompanying greater reliability of information.

The instructors were requested to indicate, on the rating scales provided, how the student impressed them relative to: (1) the student's <u>ability</u> to meet the requirements of the course, (2) the student's <u>effort</u> toward mastering the skills and materials of the course, (3) the student's general <u>attitude</u> toward the course, and other related questions. They were also to check boxes indicating their feeling as to the

adequacy of their observations upon which they based their rating. Space was provided for them to indicate instances where they had no basis for a judgment. A few instructors checked the space which indicated "no basis for judgment," but the large majority felt that they had observed the students to a sufficient extent for making their ratings.

Since the blank was first submitted at mid-term, a space was provided for the instructor to write the mid-term grade. While this probably required the instructor to refer to his records, the fact that he already had recorded mid-term scores for his department, coupled with the alphabetical arrangement of the blanks, made it relatively easy to supply this piece of information. At the end of the quarter when the blanks were again submitted, the instructors were <u>not</u> asked to indicate a grade on the blank. The end-of-quarter grades were obtained later from the instructor's final grade report.

Most of the blanks were returned within one week after being sent to the instructors and all blanks from the first set were returned before the second set was released to the staff. There was a 100% return on the first set and a 94% return on the second set by instructors of the Departments of Communication Skills and Natural Science.

The "Special Information Blank" (see Appendix) was introduced by the paragraph:

> Your Basic College Research Committee, at the request of Dean Erickson, is now engaged in a study of student achievement. It is of utmost importance to this study that you make as accurate appraisal of (student's name) as you can at this time. Please return this rating when completed to the Head of your Department.

Class size had been limited to permit the instructors to become acquainted with their students - to personalize the instruction. The feeling that their Dean and Department Head might expect them to know their students could have kept some instructors from using the "out" provided them and forced them to make a rating for the student where they would have checked the space indicating "no basis for a judgment" had the information blank come through different channels.

There is no way that this bias can be estimated. However, there is no reason to believe that a false report would be submitted concerning any student but only that the instructor might not be as confident of his rating as he indicated.

Since the instructors were not informed that they would be asked to make two appraisals of the students, there is no reason to believe that a special effort was made to make the two ratings agree. The fact that they were forced to evaluate the particular students of the study on the specific items of the questionnaire and did not have to do so for

students generally might have influenced their second rating - tending to make it conform to the first rating. This factor could not be avoided without great expense and waste of faculty time.

There was some previous knowledge on the part of the instructional staff as to the general nature of the study - they knew, at least, that it was to be a study of low-ability students. An effort was made to eliminate any bias this prior knowledge might produce. The questionnaire included questions that could refer to high ability students as readily as to low-ability students. For instance: "How often has this student done more than just the work that was assigned?" The questions important to this study were also worded in such a way as to apply equally well to high-ability as to low-ability students.

An additional precaution was taken to eliminate this type of bias. The packet of questionnaires for each section of each course not only included blanks for low-ability students (those who received one of the lower four standard scores on the A.C.E. Psychological Examination), but also included approximately an equal number for high-ability students (those who received one of the upper four standard scores on the same test). Since all of the blanks were arranged alphabetically in each section, there was no systematic arrangement of the low-ability students in any

of the various packs. The responses by the instructors gave no evidence that they considered the study to be concerned only with low-ability students.

A questionnaire (see Appendix) seeking information about the students' general interest in Basic 111 and Basic 121 and about the amount of their studying for those classes was sent with explanation and instructions to the students themselves. These questionnaires were presented to the students by their Basic 111 instructors during the last week of class in fall quarter, 1952. As was stated in the instructions, the completed questionnaire was to be sealed in an addressed envelope that accompanied the questionnaire. The sealed letter was to be handed to the Basic 111 instructor who would place it in the mail. By this procedure a better than 91% return was achieved. Absences and no response by the students accounted for most of the remaining questionnaires.

Since this questionnaire to the students was being handled by the Basic 111 instructors at about the same time as the second report by instructors on the "Special Information Blank," it was necessary to question all students about whom the instructors would be submitting information. Had only the low-ability group been questioned, it would have informed the instructors as to the group of students of greatest interest to the study. Such information might have

influenced their second response to the "Special Information Blank."

There was apparently some misunderstanding on the part of some students in connection with one or two of the questions. The introductory statement included the following: "To get an over-all picture, it is necessary to question a few students that are representative of the general student body. You happen to be one of the students selected to represent your group. An accurate response to the questions below will be greatly appreciated by the Research Committee." Some students apparently interpreted this to mean that they were to represent their Basic 111 group. They responded to the questions relative to Basic 111, but did nothing with questions pertaining to Basic 121.

Another weakness in the instrument was that it allowed only four choices instead of five in describing the students' studying. They were given the following choice of response:

I studied { ( ) a little ( ) some ( ) quite a bit ( ) very hard } earlier in the quarter.

Since there was no mid-position, the students' natural tendency to show themselves in the best light possible was not only permitted, it was encouraged. Had five choices been allowed, many students might have indicated a medium amount of studying.

In trying to obtain information about the students' study effort from the dormitory assistants, a "Study Attributes Check List" (see Appendix) was submitted to the assistants through the residence counselors. There is a real question as to the extent to which the dormitory assistants were able to observe study effort on the part of the students. Each assistant was assigned to a precinct of from 40 to 80 students. Not only was that too many students for the dormitory assistant to know intimately in the space of one quarter, but not all students did their studying in the dormitory.

The check list was submitted on Monday following Thanksgiving Recess - this was thought to be late enough in the quarter to allow the assistants to become somewhat acquainted with their charges, but not too far into the quarter that their judgment would be influenced by end-of-term studying.

It was requested that the assistants complete the blanks on a basis of the information about the student that they had at that particular time. They were not to make special inquiry or start an investigation since an evaluation of a term's work was desired - not a week's report. To what extent this suggestion was followed is not known.

Many dormitory assistants indicated that they had no basis for a judgment at that time. Some indicated that they were more or less guessing on some of the opinions since it

was difficult for them to be certain. Even though they might have found the student at his desk whenever they entered his room it did not necessarily indicate that he was studying.

Information about the student's interest was obtained from the students themselves. Information about the student's apparent ability and his class attitude was obtained from the instructors. However, information about the student's study effort came from three sources - the instructors, the students, and the dormitory assistants.

Three different types of instruments were used to obtain information about the student's study effort from the respective sources. They were:

To the Basic instructors:

Please indicate how this student impresses you by placing a check (✓) mark at the proper place along the scale lines:

Student's effort toward mastering the skills and materials of this course:

Not Trying Doing a Fair Amount Working very Hard

To the Students:

Check (✓) the words in the sentences below that best describe your studying for your Basic courses so far this quarter:

I studied { ( ) a little ; ( ) some ; ( ) quite a bit ; ( ) very hard } earlier this quarter.

I am studying { ( ) less ; ( ) the same ; ( ) more } now.

To the dormitory assistant:

Please check (✓) those expressions which best answer the question as far as the above named student is concerned:

To what extent has this student been studying?

Not at all ( ) Very little ( )
About as much as most students do ( )
Quite a bit ( ) Very hard ( )

The fact that these instruments are not the same might mean that they are not measuring the same variable. However, they are all dealing with effort of some sort. There is a possibility that the instructors did not estimate the studying the student was doing since the scale did not request that specifically. Be that as it may, the information from the students and the dormitory assistants is certainly related to the effort the student is making toward mastering the skills and materials of the courses.

The information from the dormitory assistants was of a general nature in that they could not tell whether a student was studying Basic 111 or some other course. All that they could hope to indicate was whether the student studied at all

in the dormitory.

Not only is the variable "Effort" multi-dimensional from the viewpoint of any one observer in his situation, but when rated by different observers, under different conditions, and in different situations it is very doubtful whether it is the same variable that is being rated. However, there might be sufficient similarity between ratings by observers to indicate that they are seeing different aspects of the same thing. That was the hope of this study.

CHAPTER V

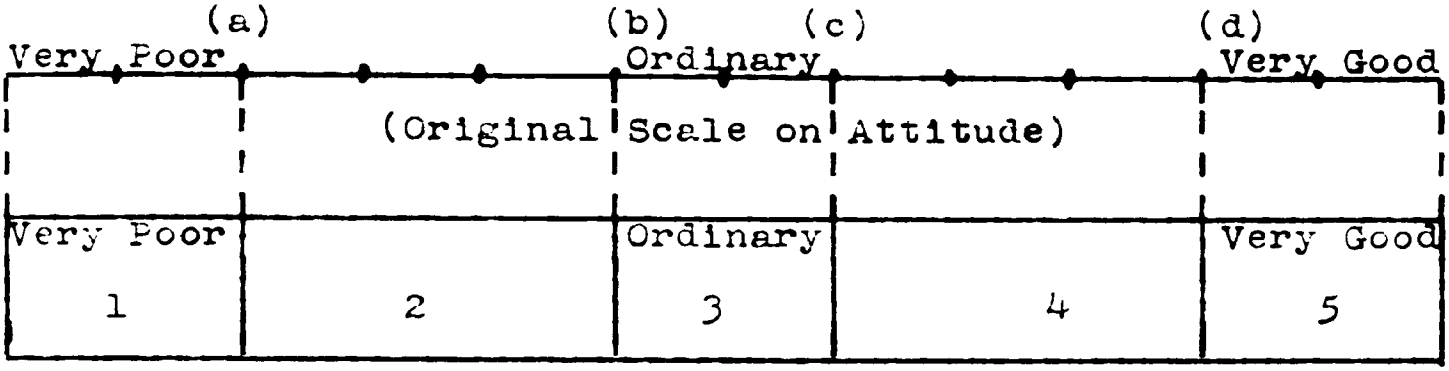
METHODS OF SUMMARIZING AND ANALYZING THE DATA

The rating scales used in this study were not absolute scales nor were they standardized instruments. As has been pointed out earlier in this dissertation, the variables concerned are multi-dimensional. One observer might not rate a student the same as another observer. Or, the same observer might not give the same rating at two different times.

The scales on the "Special Information Blank" that was sent to the Basic 111 and Basic 121 instructors were subdivided into approximately ten units by cross marks on the linear scale. Word definitions of the ratings appeared at either end and in the center of the scale. For instance, the scale used for obtaining information about the student's general attitude toward the course was as follows:

Very Poor Ordinary Very Good

Since the ratings probably were not as accurate as the ten subdivisions indicate, it was decided to summarize the ratings by subdividing the scale into five intervals in the manner shown on the next page.



(Attitude Scale Having Five Subdivisions)

Each of the five intervals was given an index number to identify and differentiate it. All ratings falling along the scale in a space corresponding to one of the five intervals were given that interval's index number regardless of where they fell in the interval. Ratings falling on points (a), (b), (c), or (d) (shown above), the points between consecutive intervals, were included in the interval containing the word definition of the rating. That is, ratings falling on point (a) were placed in the first interval and given the index 1, ratings falling on points (b) or (c) were placed in the third interval and given an index 3, ratings falling on point (d) were included in interval five and given the index 5.

It might be argued that since the subdivisions are of unequal size, the distribution of ratings is thereby distorted in grouping. However, it is felt that in using a rating scale where parts of the scale are defined while other parts are not, the person doing the rating might be drawn

toward the defined portions. The observer who feels that the individual being observed is about ordinary might be drawn to the word "ordinary" on the scale. If he checks between "very poor" and "ordinary" he probably is certain that it should not be "very poor" nor "ordinary," but somewhere between the two.

If an instructor feels that his student has a poor attitude toward the course - where might he check along the scale? Certainly not "very poor," the student is thought to be better than that. Will he check "ordinary"? Not if he considers "ordinary" to mean average - approximately halfway between "very poor" and "very good." Poor will probably be somewhere between "very poor" and "ordinary," but might not be halfway between.

Since the ratings probably are not as precise as the original scale indicated, the grouping into five intervals as indicated will probably not distort the distribution appreciably.

The replies on the check list, sent to the dormitory assistants, relative to the student's studying were forced into five categories by the nature of the check list. These categories had the index numbers one through five assigned to them in order of increasing study. Index number 1 was assigned to the statement "Not at all" and 5 was assigned to the statement "Very hard." While the check list appears

to be definitive, it is as indefinite as the rating scale sent to the instructors. For instance, the statement: "About as much as most students do." - how much is that? The amount that this indicates will undoubtedly vary from observer to observer. However, its position in the list of choices indicated that it was centrally located between little studying and much studying. Perhaps the dormitory assistants realized that. At best, it was a crude measure.

In trying to get an evaluation of the studying done by the student, two questions were asked instead of one. This allowed the student to somewhat describe the nature of his studying during the quarter. The request for information was worded in the following manner:

> Check (✓) the words in the sentences below that best describe your studying for your Basic courses so far this quarter:

I studied: { ( ) a little ; ( ) some ; ( ) quite a bit ; ( ) very hard } earlier in the quarter.

I am studying: { ( ) less ; ( ) the same ; ( ) more } now.

It was felt that this set of statements might partially overcome the tendency to report the best studying done during the quarter as being typical of the entire quarter's work. This type of response, however, made it somewhat

difficult to rank the responses into an ascending scale of study effort.

However, the possible combinations of response were listed and grouped into seven categories - each of which seemed to involve a better study pattern than the previous one. The categories are listed as follows:

Index Number	Type of Studying Reported
1	A little studying earlier - less now.
2	A little studying earlier - the same now, or Some studying earlier - less now.
3	A little studying earlier - more now, or Quite a bit of studying earlier - less now.
4	Some studying earlier - the same now, or Studied very hard earlier - less now.
5	Some studying earlier - more now, or Quite a bit of studying earlier - the same now.
6	Studied very hard earlier - the same now, or Quite a bit of studying earlier - more now.
7	Studied very hard earlier - more now.

The index numbers identify the category into which the various ratings by the students were placed. Since the students rated themselves but once during the quarter, there

is no estimate of the instrument's reliability.

The instructors rated each student twice and thereby furnished an index of reliability for the instruments that they used in rating the students. Since the instructors rated the same students twice, a combined rating was obtained by adding the index numbers given the individual ratings. As each individual rating was supposed to be a summary of the quarter's performance until the time the rating was made, the second rating was based on more observations than was the first. The additional observations were made during the three - or four - week interval between the two ratings. By adding the two index numbers, the trend during the quarter was allowed to influence the combined effort rating. In most instances the two ratings by the same instructor were in agreement with each other so the trend effect was slight.

To obtain a pooled effort rating from the three separate ratings by the various sources, it was decided to weight those ratings in the following manner:

Type of Rating	Weight
Instructors' combined effort rating	2
Students' rating of own effort	2
Dormitory assistants' rating	1

It was felt that the instructors, having made two ratings, should have their judgments weighted twice that of the

dormitory assistants' single judgment. It was also felt that the students know as much about their own studying as instructors know about that studying and should therefore have their rating count as heavily as does the instructors' rating.

The pooled effort rating score was determined in the following way: The index numbers for the various effort ratings on each student were multiplied by the weighting factors mentioned above. That is, the index number identifying the study pattern described by the student himself was multiplied by <u>two</u>, the sum of index numbers identifying the effort ratings by the Basic 111 (or Basic 121) instructor was multiplied by <u>two</u>, and the index number identifying the category checked by the dormitory assistant as describing the student's studying was multiplied by <u>one</u>. These values were then added to obtain a pooled effort score for the student involved.

It was desired to express the effort score as a percentage. This was accomplished in the following manner: The possible range of effort scores was determined by computing the minimum effort score that could possibly be given and the maximum effort score that could possibly be given, then subtracting the minimum from the maximum. Each effort score was then adjusted by subtracting the minimum effort score from it to get an effective or adjusted effort score. By

dividing each adjusted score by the possible range of effort scores and multiplying by 100, it was possible to express the pooled effort rating for each student as a percentage score.

Two separate percentage scores were determined for each student - one of which included the ratings by the Basic 111 instructor and the other included ratings by the Basic 121 instructor. The first score was used for comparisons with Basic 111 grades while the second was used for comparisons with Basic 121 grades. This was done to provide as accurate a picture as possible of the effort that was put forth toward mastering the course concerned.

Two different methods were used in defining low effort - high effort. One procedure was to call those students receiving a percentage effort score of 50 or below "low-effort" and the others "high-effort." The second method was to call the individuals receiving the lowest 31 scores "low-effort." This amounted to approximately 26% of the group. Those individuals receiving the highest 31 scores were called "high-effort."

Students that were called "low-effort" under the second definition above received percentage effort scores of 45% or below if Basic 111 ratings were involved and 42% or below if Basic 121 ratings were involved. "High-effort" students, according to the second definition, received percentage

effort scores of 66⅔ or above if Basic 111 ratings were involved and 60% or above if Basic 121 ratings were involved in their derivation.

The second method of defining low and high effort leads to more positive results since comparisons are between grades given students widely separated as far as pooled effort ratings are concerned. By ignoring the middle half of the total group, greater confidence can be placed upon the titles "low-effort" and "high-effort."

The ratings made on the variables under study were tabulated and the results were either placed in a single classification frequency table, where a single variable was to be examined by itself, or in a two-way classification frequency table where two variables were to be compared. For instance:

Table 5 - 1

Basic 121 Instructors' Ratings of Students' Effort to Meet Basic 121 Requirements. (Combined Ratings.)

Index	Frequency
1	1
2	3
3	11
4	18
5	37
6	16
7	17
8	3
9	5
10	0

Table 5 - 2

Basic 111 Instructors' Ratings of Students' Attitude Toward Basic 111.

First Rating \ Second Rating	1	2	3	4	5
1					
2		4	5	1	
3		6	20	7	1
4		1	7	30	4
5			2	6	15

To analyze and compare the data it was first necessary to determine the arithmetic mean and the standard deviation for each of the frequency distributions. The mean represents the index value most centrally located relative to the entire distribution, while the standard deviation indicates the spread or scatter of ratings or grades about that central value.

Distributions of the same general nature may differ in level, that is - their means, or differ in the extent to which they are dispersed about their means or both. Significance tests may be used, under proper conditions, to test differences between means or differences in variance (standard deviation squared) to determine, with a previously specified risk of rejecting the hypothesis when it should be accepted, whether the two parent populations from which the two samples were drawn might be considered as having the same characteristics (the same mean and same standard deviation).

The significance tests used in analyzing some of the data of this study involved the t and the F distributions. The underlying assumptions and the general procedures for using these statistics are as follows:

In testing the hypothesis that two parent populations have the same mean, it is first necessary to test whether they have the same variance (unless this fact is somehow

previously known).

Let us say that we have two sample groups - one group having n measurements, a mean value of $\bar{x}$, and a standard deviation s_x; the other group having m measurements, a mean value $\bar{y}$, and a standard deviation s_y. The parent population from which the first group was selected had a mean M_x, and a standard deviation σ_x; while the parent population from which the second group was selected had a mean M_y, and a standard deviation σ_y. These parameters (M_x, σ_x, M_y, σ_y) have values that are not known. The only clues we have as to their possible values are furnished by the sample statistics $\bar{X}$, S_x, $\bar{Y}$, S_y.

The first hypothesis to be tested is whether the two parent populations have equal variances, that is: $\sigma_x^2 = \sigma_y^2$. At this point, we should state the risk that we are willing to take of rejecting the hypothesis when actually it should be accepted. This is called Type I error.

A rather common risk assumed in the field of educational research for Type I error is 5%. Such a risk is assumed throughout this study.

For testing the hypothesis $\sigma_x^2 = \sigma_y^2$ we use the F statistic which is defined as $F = \frac{S_x^2}{S_y^2}$, the ratio of the variances for the two samples. The assumptions underlying the use of this statistic are: The observations are to be random samples from normal populations and the hypothesis is assumed to be true. Under these assumptions, the sampling distribu-

tion for the F statistic has been computed and may be found in tabular form in some elementary statistics textbooks or books of statistical tables.[1]

Rejection regions may be determined by obtaining F values from the table or computed from them. The F values used must take into account the number of observations in the two samples and the size of risk assumed for Type I error. If the value for F that is computed by dividing S_x^2 by S_y^2 is included in either rejection region, the hypothesis is rejected and the parent populations are said to have significantly different variances. If, however, as happened in every instance in this study, the computed F value does not fall within one of the rejection regions, then the hypothesis is accepted and the parent populations are said to have the same variance.

Once the hypothesis of equal variances is accepted, a new hypothesis is proposed for test - namely: $M_x = M_y$, that is - the means of the two parent populations are equal. Accepting the same risk as before of rejecting the hypothesis when it should be accepted, we operate at the 5% level of significance.

It has been accepted that the variances of the parent

---

[1] One such set of tables may be found in Dixon, Wilfred J. and Massey, Frank J. Jr., Introduction to Statistical Analysis, McGraw-Hill, 1951, pp. 310-313.

populations are equal (by the previous consideration) but the value of that variance is not known. It will be necessary to estimate that value as best we can, using the known variances of the two samples. This we do by using the formula: $S_p^2 = \frac{nS_x^2 + mS_y^2}{n + m - 2}$, where S_p^2 is the pooled mean square estimate of σ^2.

The statistic that we use in testing the hypothesis $M_x = M_y$ is the t statistic which is defined as:

$$t = \frac{\bar{X} - \bar{Y}}{S_p\sqrt{\frac{1}{n} + \frac{1}{m}}}$$

The assumptions underlying the use of this statistic are as follows: Both populations are to have normal distributions with the same variance. Under these assumptions, the distribution for the t statistic has been computed and placed, in tabular form, in various statistics texts.[2]

By taking into account the number of observations in the two samples and the level of significance, it is possible to obtain critical t values from the tables and determine rejection regions.

When the value for t is computed using the information from the two known samples, it may then be seen whether that value falls within one of the rejection regions. If it does,

---

[2] Ibid., p. 307.

the hypothesis ($M_x = M_y$) is rejected and it is concluded that the two populations from which the samples were drawn have significantly different mean values.

If, however, the t value, as computed, does not fall in either rejection region then the hypothesis is accepted that the two mean values are the same. That is, there is no significant difference between the means of the two parent populations.

Both tests of significance described above are based upon the assumption that the samples have been drawn from normal populations. Is there reason to believe that this assumption is justified for the populations from which the samples used in this study were drawn? Since there is no prior knowledge of the nature of these populations, the question can only be answered on a basis of the samples and of the factors that influenced the nature of the samples.

Some of the samples were composed of scores received on the final examinations in Basic 111 or Basic 121. The students to whom these scores in the sample were given had been picked by means of the A.C.E. Psychological Examination scores as being poor students and the students had been placed in a low-effort group (or in a high-effort group) by the technique previously described.

Were the A.C.E. Psychological Examination and the effort rating technique mentioned previously perfectly discriminat-

ing, there might be reason to expect the distributions of final examination scores to be skewed. The low-ability - low-effort students should receive very low grades with the largest number getting failing grades and a few getting passing grades. And the low-ability - high-effort students should have grades mostly above the fail line with very few below it.

However, since neither the A.C.E. Psychological Examination nor the effort rating technique is anywhere near perfect, it was possible for students who learned quickly with little apparent study effort to be classified as low-effort students. The presence of these higher-ability students in the low-ability group would tend to mean additional passing grades. These additional grades would probably cluster just above the passing point with some (decreasing in number with increasing grade) scattered up the grading scale. The two groups being combined would tend to show a nearly normal grade distribution. This leads to the belief that the populations from which the samples in this study came were normally distributed. The appearance of the sample distributions in this study lend credence to this belief. At least, there is no evidence that the populations were not normal.

It must be pointed out that the parent populations from which the samples of this study were drawn were not

composed of students, but were distributions of grades (or ratings) that might reasonably have been expected to have been given to a certain group of students. The group in question was composed of all possible persons who might have entered M.S.C. fall quarter, 1952 to start their first quarter of college work in the normal course of events, and who would have received one of the lower four standard score measures on the test-retest of the 1949 edition of the A.C.E. Psychological Examination and who would have, in the natural course of events, been rated as low-effort by the effort rating device used in this study. That distribution of grades was assumed to be normally distributed and the particular sample that was obtained in this study was assumed to have been randomly drawn from that parent population of Basic 111 final examination grades.

The final assumption of random selection is not so easy to defend since the sample of the study was selected by the convenient method of choosing only grades given to students who actually enrolled and completed their first quarter's work at Michigan State College. Why other possible applicants did not come to M.S.C. is left to conjecture. There are many factors such as financial ability, geographic location, personal preferences, etc., that undoubtedly entered into the consideration.

Although the randomness of the selection is question-

able, nothing related to the study might have conditioned the decision to come or not come to college unless some candidates, upon taking the A.C.E. Psychological Examination in the summer months preceding fall term, 1952 decided against coming to M.S.C. because of a feeling that they did poorly and might not be fitted for college work.

In determining the relationships between the various factors and the grades given in Basic 111 or Basic 121, the Pearson product moment coefficient of correlation was used.

Whenever two variables, that is, ratings or grades, pertained to the same student, these paired values were first indicated on a scatter diagram and then summarized in a two-way classification frequency table similar to Table 5 - 2. From such a table it was possible to determine the frequency distributions for the two variables concerned, with their means and standard deviations ($\bar{X}$, S_x and $\bar{Y}$, S_y). It was also possible to multiply the two paired values for each student and to determine the average $\left(\frac{\Sigma XY}{N}\right)$ of such products for the entire number (N) of students in the sample. The Pearson product moment coefficient of correlation was then computed by using the formula:

$$r = \frac{\frac{\Sigma XY}{N} - \bar{X}\bar{Y}}{S_x S_y}$$

Assumptions of randomness in sampling and normality

of parent populations are not assumed in the derivation of the coefficient of correlation. Any set of paired numbers, however obtained and from whatever source, has a coefficient of correlation.

While the formulation itself implies no assumptions, the coefficient of correlation is interpreted by means of assumptions that are made about the natures of the variables being compared. These assumptions are usually based upon logical considerations or intuition founded upon prior observation of the two variables. There are reasons for believing that the two factors are related prior to the computation of a coefficient of correlation - the correlation coefficient then furnishes an estimate of the extent of that relationship.

A high correlation between the variables does not of itself mean that there is a strong cause-effect relationship between the variables. However, when the probability of cause-effect has been established, the coefficient of correlation may be used to estimate the strength of the relationship.

On the other hand, if the coefficient of correlation for two variables is zero, there is no cause-effect relationship between those two variables. This is so whether a cause-effect relationship was assumed <u>a priori</u> or not. The variables are unrelated, have no common elements.

With the procedures and tools used for this study in mind, it will be possible to interpret the findings - taking care to note where basic assumptions inherent in the tools are only partially satisfied by the procedures used for obtaining the data.

CHAPTER VI

GENERAL FINDINGS AND RELATIONSHIPS

A definite effort was made at the start of this study to accurately predict which students might be of low academic ability. The extent to which that was accomplished may be determined by examining Table 6 - 1, on page 61.

It can be seen that a considerable percentage of low-ability students received a grade of C or better, not only from their instructors, but on a basis of their final examinations as well. So many were passed, in fact, that only a very small percentage were failed by either their instructors or the final examinations. Even when the instructors' mark and the final examination grade were combined, only 4.3% of the "low-ability" students received the grade of F in Basic 111, while 9.3% received that grade in Basic 121.

It may be noted that while most of the grade distributions for Basic 121 conform closely to the suggested pattern: A 7-11 B 25-29 C 45-49 D 12-16 F 0-5, the Basic 111 instructors' grades for all freshmen had an appreciably higher percentage of B's and a much smaller percentage of D's. This general tendency by the Basic 111 instructors to give the higher grades undoubtedly influenced the grades they gave the "low-ability" students, and possibly influenced the various ratings they made on those students. Distributions of ratings by the Basic 111 instructors were generally

Table 6 - 1

Percentage Distributions of Grades Received by Low-Ability Students and by All Freshmen in Basic 111 and Basic 121 During Fall Quarter, 1952.

Grade	Instructors' Grade Distribution Percentages				Final Examination Grade Distribution Percentages				Course Grade Distribution Percentages			
	Low-Ability		All Freshmen		Low-Ability		All Freshmen		Low-Ability		All Freshmen	
	111	121	111	121	111	121	111	121	111	121	111	121
A	2.5		8.1	8.4			7.6	7.9			7.9	8.4
B	25.2	8.5	40.3	25.6	2.6	6.8	23.8	25.6	8.7	9.3	35.8	28.5
C	58.0	44.0	44.1	46.4	31.2	44.1	49.8	51.3	42.6	32.2	41.9	41.1
D	12.6	33.9	5.9	14.7	47.0	40.6	15.8	12.2	44.4	49.2	12.4	17.7
F	1.7	13.6	1.1	4.3	19.2	8.5	2.9	3.0	4.3	9.3	1.5	3.8

The percentages for the "All Freshmen" distributions do not necessarily total 100% for the table shown here. A few miscellaneous grades such as the Incomplete were given, but not shown in the distribution above.

Data for the "All Freshmen" distributions may be found in Summary of Basic College Course Grades. M.S.C. Board of Examiners. January 21, 1953.

skewed toward the more favorable ratings.

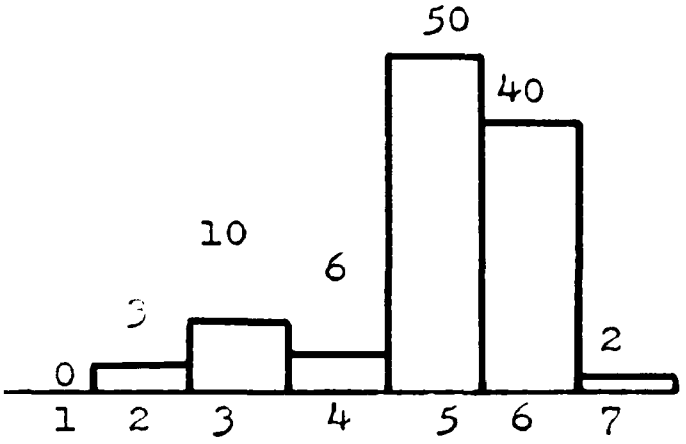
It can be seen from the table that the statements made earlier relative to the A.C.E. Psychological Examination and its predictive validity are in agreement with the facts. As a group the "low-ability" students do receive lower grades than do freshmen generally; however, there are many capable students included in the low-ability group.

In trying to determine a pass-fail point on a basis of effort, it was decided to separate the low-ability group into two sub-groups, a low-ability - low-effort group and a low-ability - high-effort group. It was felt that if the hypothesis proposed by the Basic College Educational Research Committee was correct, the mean grade given the high-effort group would be significantly higher than the mean grade given the low-effort group. The fact that the low-ability group contained many capable students made it less likely that such a separation of mean grades would be found. If some students in the low-effort groups were actually <u>capable</u> students, they would tend to get higher scores than would students of truly low ability while making the same effort.

The next factors that must be considered in understanding the problem of establishing a pass-fail point are the various effort ratings that were obtained. These findings are as follows:

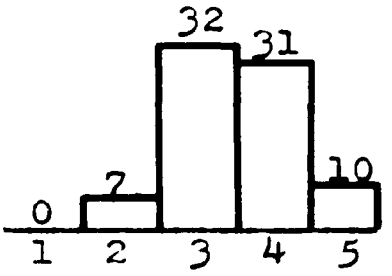
Histogram 1

Frequency Distribution of Ratings by Low-Ability Students on their own study Effort.



Histogram 2

Frequency Distributions of Ratings by Dormitory Assistants on Study Effort of Low-Ability Students.



(Index numbers increase in the order of increasing study)

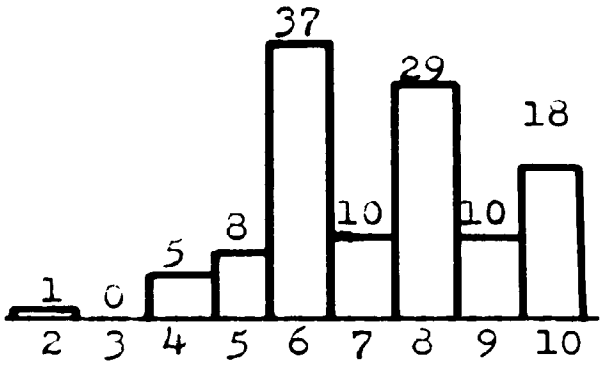
Histogram 1 shows a skewed distribution due to the natural tendency on the part of the students to try to look their best and the fact that they could not indicate a medium amount of studying on the instrument given them.

The ratings by the dormitory assistants, however, are surprisingly symmetric. Index number 1 was assigned to the "not at all" category. Apparently it was felt that all of the students in the sample group did <u>some</u> studying.

The following histograms will show the distributions of ratings on effort given by the Basic 111 and Basic 121 instructors:

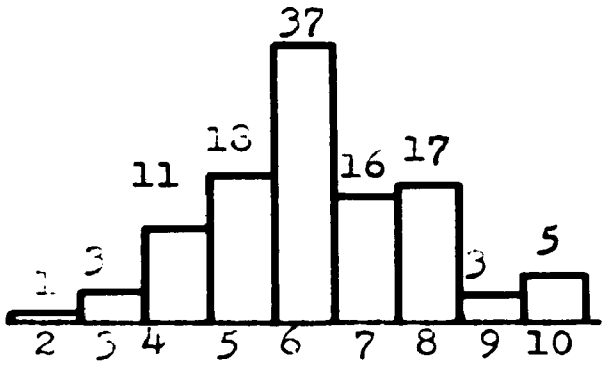
Histogram 3

Frequency Distribution of Effort Ratings on Low-Ability Students by Their Basic 111 Instructors. (Combined Ratings)



Histogram 4

Frequency Distribution of Effort Ratings on Low-Ability Students by Their Basic 121 Instructors. (Combined Ratings)



(Index number totals increasing with increasing appraised effort)

The instructors rated each student twice and each rating was assigned an index number 1, 2, 3, 4, or 5 as previously described. The index numbers for the two ratings on a single student were added to obtain an index number total.

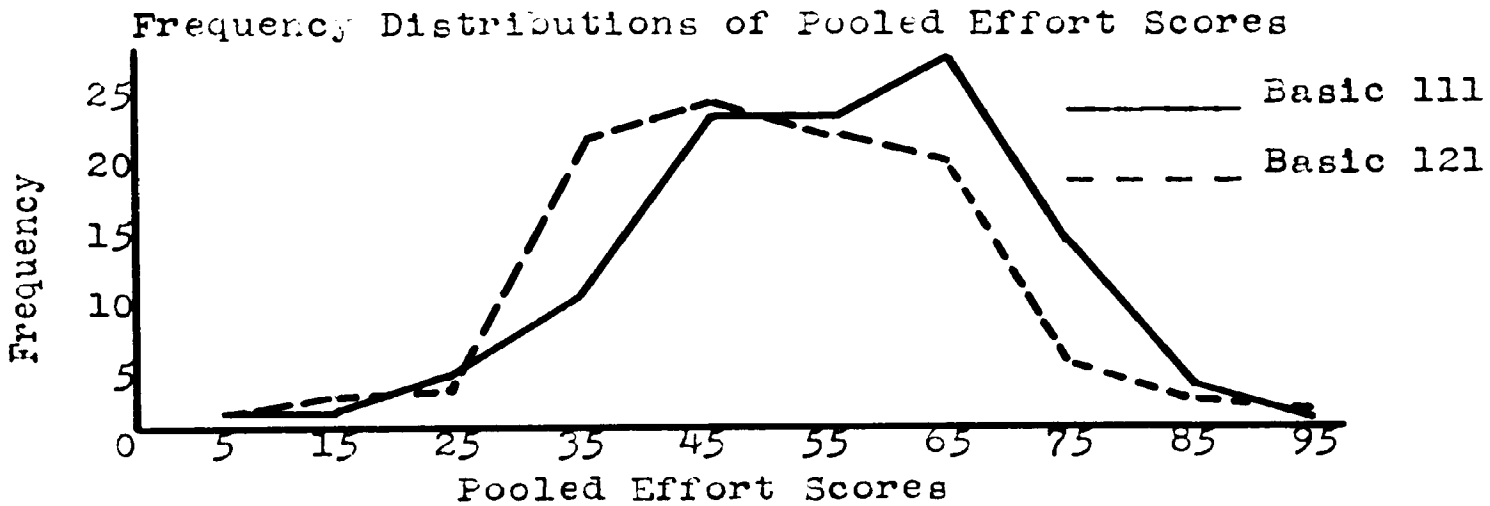
The distribution of combined ratings by the Basic 121 instructors approximates a normal distribution while that by the Basic 111 instructors is definitely skewed toward the higher (more favorable) ratings. These distributions show the same general pattern as the grade distributions for all freshmen in Table 6 - 1.

The tendency of the Basic 111 instructors to give the

students a favorable rating is similar to the tendency of the students themselves (Histogram 1). However, the instrument sent to the students forced them away from a medium rating, as was explained before, whereas the instrument sent to the Basic 111 instructors allowed unlimited choice. It appears that there is a general tendency on the part of the Basic 111 instructors to give their students favorable ratings and grades. This tendency diminishes the effectiveness of the A.C.E. Psychological Examination as a predictive instrument and decreases the discrimination of low-effort - high-effort classifications determined for this study. Such a situation reduces the chance of establishing a pass-fail line on a basis of the A.C.E. Psychological Examination scores and pooled effort ratings.

The fact that the pooled effort ratings reflect the bias shown in the Basic 111 instructors' ratings of effort may be seen in the following graph (which also shows other bias as described later).

Graph I



It might also be noticed that the students' skewed ratings when combined with the symmetric ratings by the dormitory assistants and the fairly normal ratings by the Basic 121 instructors have a tendency to give the same type of skew to the pooled effort score distribution for Basic 121 as was exhibited by the student rating distribution itself.

When the student skewed ratings are combined with the Basic 111 instructor ratings, skewed in the same direction, and the symmetric ratings by the dormitory assistants, the resulting distribution is strongly skewed in the same direction, as is shown in the graph.

To answer the question that was of concern to the Basic College Educational Research Committee, namely: Is it possible to establish a pass-fail point for grades given in the Basic College on a basis of the study effort made by low-ability students?, it was necessary to compare the grades received by the low-ability students with their pooled effort rating scores. This was done first by examining the two-way frequency tables and then by testing the significance of the differences between the mean scores received by the low and high-effort groups.

The answer in every case was the same - no such pass-fail point could be established on a basis of the data obtained in this study.

Such a conclusion may be confirmed by examining the following two-way classification frequency distribution tables comparing pooled effort rating scores with course, instructor, and final examination grades given in Basic 111 and Basic 121:

Table 6 - 2

Pooled Effort Rating Scores and Course Grades Received in Basic 111 and Basic 121

Basic 111

Course Grade	0	15	29	43	57	71	85	100
A								
B				2	3	3		
C		2	5	16	15	9	1	
D	1	3	9	11	17	5	2	
F		1		3	1			

Pooled Effort Scores (Percent)

r = .22

Basic 121

Course Grade	0	15	29	43	57	71	85	100
A								
B			1	5	1	2		
C			5	18	6	2	2	
D	2	3	17	14	16	3	1	
F		2	1	4		1		

Pooled Effort Scores (Percent)

r = .17

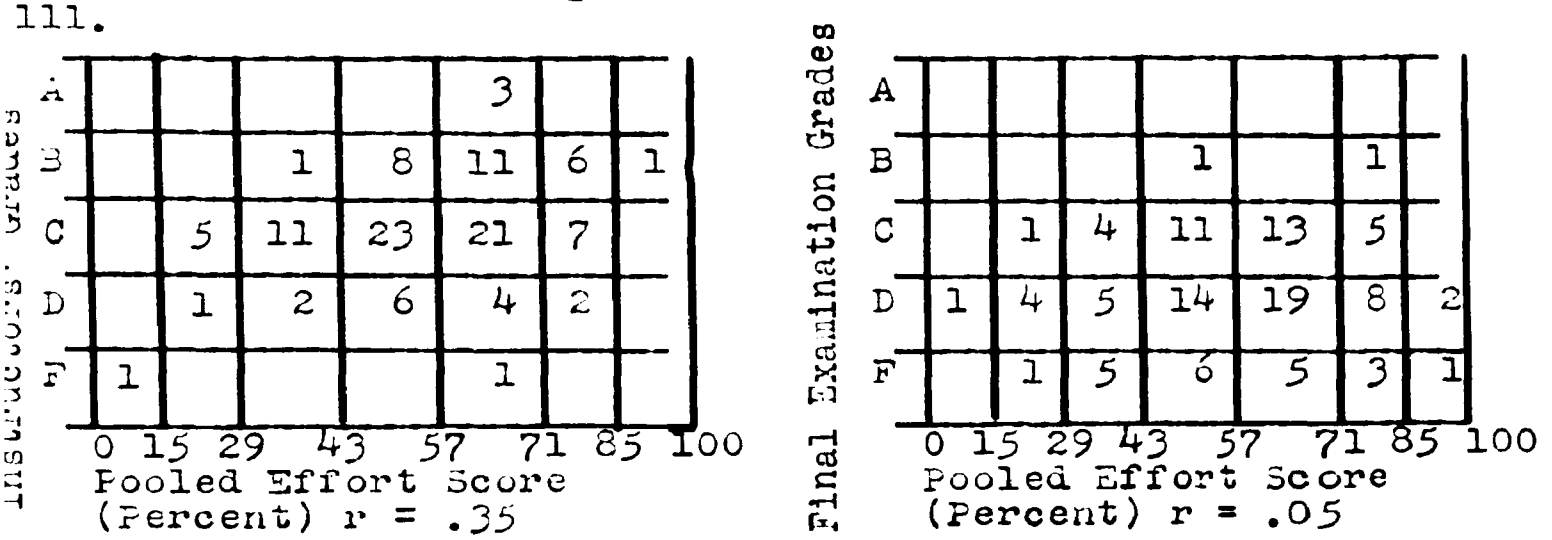
It can be seen that the pooled effort ratings are not related sufficiently to the course grades to be useful in determining F grades. For instance: If it were decided that students in Basic 111 who received an effort rating of 29% or below should be failed and all other students who received comparable course grades (instructors' score ≠ final exam score) as those received by low-effort students should also fail, then it can be seen from the table that all but eight students in this entire group would fail -

even though the main group received higher pooled effort ratings. This contradiction cannot be resolved by choosing some other effort score as a criterion.

This same condition is true with instructors' grades and final examination grades.

Table 6 - 3

Basic 111 Effort Rating Scores and Grades Received in Basic 111.



The Basic 111 instructors gave but two F grades to members of the low-ability group - one in the lower half of the pooled effort score range, the other in the upper half. While the Basic 111 final examination placed more students in the F category, they were spread over the effort scale to such an extent that there seemed little relationship between the grade F and the pooled effort scores. Certainly the table above does not indicate a pass-fail point for Basic 111 on a basis of the pooled effort rating scores.

The corresponding table for Basic 121 presents essentially the same pattern:

Table 6 - 4

Basic 121 Effort Rating Scores and Grades Received in Basic 121.

Instructor's Grade

A							
B				6	1	1	
C	1	1	8	22	9	4	2
D	1	3	14	8	9	2	1
F		1	2	6	2	2	

0 15 29 43 57 71 85 100
Pooled Effort Score
(Percent) r = .17

Final Examination Grade

A							
B			2	4	1	1	
C			12	19	8	5	2
D	2	3	8	14	13	3	1
F		2	2	4			

0 15 29 43 57 71 85 100
Pooled Effort Scores
(Percent) r = .16

Again the grades of F are dispersed over a range of pooled effort scores as are the passing grades. No pass-fail point is apparent in the distributions above.

As further evidence that the conclusion from these observations is valid, the statistical tests of the significance of the difference between mean grades for parent populations from which the "low-effort" group and "high-effort" group were drawn, show no significant difference at the 5% level. This is true under either of the two definitions for "low-effort" or "high-effort" described in chapter five. The second definition, it might be recalled, considered the 31 individuals receiving low-

est pooled effort ratings as having grades in the "low-effort " group while grades given the 31 individuals receiving highest pooled effort ratings constituted the "high-effort" group. Even grades given students who have received such diverse pooled effort ratings cannot be judged as being significantly different according to the tests used.

Table 6 - 5

Mean Scores and Standard Deviations for Groups of Low-Ability Students Classified as Low-Effort and High-Effort (Instructors' Scores and Final Examination Scores were based on a 15 point Scale. Course Grade Totals range from 2 to 30)

Type of Score	Low-Effort Group 31 individuals receiving lowest pooled effort rating scores		High-Effort Group 31 individuals receiving highest pooled effort rating scores	
	Mean	S.D.	Mean	S.D.
Basic 111 Instructors' Score	7.71	1.73	9.45	2.04
Final Examination Score	4.84	2.23	5.29	2.13
Course Grade Total	12.72	3.03	14.59	3.29
Basic 121 Instructors' Score	5.55	1.77	6.42	2.10
Final Examination Score	5.84	2.06	6.32	1.95
Course Grade Total	11.39	3.20	12.81	3.64

This table shows that in every instance the high-effort group received a higher mean score; however, the differences are not sufficient to be statistically significant at the 5% level. It might be noted that the mean course grade score received by the low-effort group in Basic 111 would be given the letter grade of D while the corresponding mean for the high-effort group would receive a grade of C. In Basic 121, mean course grades received by both low and high-effort groups would have received the D grade.

The fact that mean scores for the high-effort group are consistantly higher than corresponding mean scores for the low-effort group indicates that effort was, to a small extent, being reflected in the marks given in Basic 111 and Basic 121. Effort was not related sufficiently for using it as a criterion by which to make pass-fail judgments, however.

In Chapter five, care was taken to point out the underlying assumptions for the tests of significance that were used in analyzing the data. It was recognized that those assumptions were not fully satisfied by the procedures for collecting data. In spite of this, the tests determined the same conclusion that was reached independently by the Basic College Educational Research Committee

from study of the scatter diagrams upon which tables 6 - 2, 6 - 3, 6 - 4, and 6 - 5 are based.

The original hypothesis of the Basic College Educational Research Committee that "among a group of students, all low in ability, effort should be a significant factor differentiating those who pass from those who fail" is not necessarily refuted by the inability of this study to establish a pass-fail point. As has already been pointed out, the study was not able to determine a group of students who were "all low in ability". It is also probable that the pooled effort ratings were not accurately estimating the effective study effort being made by the students. An attempt was made to get as complete an estimate as possible by recording information and observations from several observers - each seeing the students' study effort from a different viewpoint. However, by studying the ratings from the various observers, it becomes obvious that they are not rating the same variable. For example:

Table 6 - 6

Correlations of Effort Ratings by the Various Observers on the Same Low-Ability Students

Observer Classification	1.	2.	3.	4.
1. The Students Themselves		.02	.16	.24
2. The Dormitory Assistants			.27	.22
3. The Basic 111 Instructors				.39
4. The Basic 121 Instructors				

It is not surprising that there was so little agreement between the various types of ratings. They were obtained from different observers having differing amounts of experience in making objective evaluations of student effort, the observations were made in different environmental situations, and the ratings were recorded on different, non-standardized measuring instruments. No attempt was made, prior to collecting data, to obtain agreement among the various observers as to what constituted study effort.

The self-evaluations of the students in the study have, as a general rule, lower correlations with other effort ratings than do estimates by any of the other types of observers. This is understandable in light of their lack of experience in comparing various study efforts[1] and in light of the personal subjectivity involved in judging their own performance.

The most agreement occurred between ratings by the Basic 111 instructors and those by the Basic 121 instructors. There were several factors that would foster agreement, such as: Both sets of ratings were recorded on the same measuring instrument. Both sets of observations

---

[1] In connection with this it is interesting to note that when the student ratings of their own study effort were correlated with the number of hours per week they listed as having used for studying their Basic 111, the correlation was only .32. For Basic 121 it was .35.

were made in comparable classroom environments. Both sets of judgments were made by experienced teachers, persons well acquainted with the results of student effort. But, even though these factors favored similarity of rating, the extent of that similarity was surprisingly small (.39).

When all of the factors promoting divergence among the various effort ratings are considered, it is not too difficult to understand why the correlations listed in the table were so low. It can also be seen why the pooled effort rating as determined in this study might not represent the actual study effort put forth by the student. The student made the only estimate of that total effort and that estimate showed a definite bias; the other ratings were but estimates of partial study effort, as observed in restricted environments.

Even though the pooled effort rating was but a first approximation to a perfect representation of the students' effort, it did reveal some evidence that the hypothesis proposed by the Basic College Educational Research Committee might be justified; however, it was too crude an approximation to be used for determining a pass-fail point.

While there was apparent disagreement between observers as to what might constitute student effort, each

individual instructor seemed to have a definite idea as to what the term meant. The correlations between the first and second ratings made by the instructors were .81 for Basic 111 and .62 for Basic 121 paired ratings. The second ratings were made several weeks after the first and were therefore based on more observations. Student effort or its overt appearance would probably not remain constant during that interval. For these reasons, the correlations are considered to be relatively high.

Even though the instructors recognize certain evidences of effort being exhibited by their students, their appraisal of effort is not reflected to any appreciable extent in the grades that they give. The correlation between instructor's rating of student's effort and instructor's mark was only .38 for Basic 111 and .22 for Basic 121. Since these correlations were so low, it is questionable that the instructor's appraisal of effort could be used as a basis for establishing a pass-fail point. The distributions are shown in the following table:

Table 6 - 7

Frequency Distributions for Instructors' Ratings of Students' Efforts and Instructors' Grade

Basic 111

Instructors' Rating of Effort	F	D	C	B	A
10	1	1	7	9	1
9			6	3	1
8		5	12	10	1
7			9	1	
6		4	23	6	
5		2	6		
4		3	1	1	
3					
2	1				

Instructors' Grade

r = .38

Basic 121

Instructors' Rating of Effort	F	D	C	B	A
10		2	2	1	
9		2	1		
8	2	2	11	2	
7	4	3	7	2	
6	2	13	19	3	
5	1	9	7	1	
4	3	5	3		
3					
2	1	2		1	

Instructors' Grade

r = .22

Here again it is not possible to establish a pass-fail point, as was surmised from the correlations. The F grades are scattered throughout the various effort ratings.

Apparently "Effort" as it is recognized and observed by the instructors is not influencing the instructors' grading to any appreciable extent, and "Effort" as it is indicated by the pooled effort ratings, while slightly related to achievement, is not sufficiently discriminating to be used for the establishment of a pass-fail point.

In investigating how instructors' marks might "meaningfully supplement" final examination grades, factors other than effort were considered. These factors were the student's general attitude toward the course as appraised by the instructors, the instructor's judgment as to the student's ability to master the skills and materials of the course, and the student's stated interest in the course.

It was felt that the impressions the instructor received relative to the student's attitude and ability might influence the grade given the student by that instructor. The extent of that influence might be judged by examining the correlations of instructors' ratings with instructors' scores and with final examination scores, the latter correlation being used as an index of the relationship between the variable that was rated and the performance level of the student. Any difference between the two correlations will be examined to determine whether final examination grades have been "meaningfully supplemented" by the instructors' marks with regard to the variable under consideration.

The findings of this study relative to the instructors' appraisals of the students' attitudes are as follows:

Table 6 - 8

Correlations of Instructors' Ratings of Students' General Attitudes Toward Basic 111 and Basic 121 with Instructors' Marks and Final Examination Scores.

Type of Grade	Basic 111 Ratings	Basic 121 Ratings
Instructors' Marks	.40	.24
Final Examination Scores	.12	.07

It can be seen that the instructors' marks do reflect the instructors' appraisal of student attitude to a greater extent than do the final examination scores.

The student's general attitude toward the course should include many factors that are educationally desirable. For instance, a good healthy attitude might include a desire to learn more about the subject and how it relates to all other of life's endeavors. Such a desire might lead the student, in future years, to continue his study and eventually become a leader in the field.

It might indicate qualities such as cooperativeness, a constructively critical attitude, or other equally worthwhile factors. To the extent that the instructor's judgment of the student's general attitude does take into account factors of the type just described, it can be interpreted that the instructors' mark is "meaningfully" supplementing the final examination grade.

One of the factors that might enter into the students' general attitude and about which information has been obtained is the student's interest in the course. The correlations found in this study between the student's stated interest in the course and the instructor's rating of the same student's general attitude toward the course are -.02 for Basic 111 and .05 for Basic 121. Whatever the instructor's ratings include, they do not reflect the interest that the students claim to have in the course. If the students are really interested to the extent that they have indicated, they are not conveying an impression of that interest to their instructors in Basic 111 or Basic 121.

Apparently the instructors have largely determined their opinion of the student's general attitude by mid-term. This is indicated by the correlations between their first and second ratings which were .71 for Basic 111 ratings and .67 for Basic 121 ratings.

Although each instructor largely agrees with himself, as shown above, there is little agreement between the Basic 111 instructors as a group and Basic 121 instructors in rating the same students' general attitudes toward those respective courses. For the low-ability group, the

correlation was .39 and for the high-ability group as previously defined, it was .15. This might tend to refute the idea that students have a general attitude toward studying or studies that largely determines their attitudes toward specific courses. They certainly do not claim to have the same interest in both courses. The low-ability students' stated interest in Basic 111 correlated but .08 with their stated interest in Basic 121. For the high-ability group, the correlation was .00.

It would appear from the findings above that the instructors' rating of the students' general attitudes do relate more strongly to the instructors' mark than to the final examination grade. It apparently does not indicate the students' stated interest in the course. It is a fairly reliable measure and in-so-far as it reflects educationally desirable factors, it can be interpreted as "meaningfully" supplementing the factual and skill aspects of the final examination grades.

The A.C.E. Psychological Examination proved to be a poor predictor of the marks the students received in Basic 111 and Basic 121. The question arises: Would the instructor's own estimate of the student's ability be a better predictor? This will be considered with some re-

lated questions, such as: To what extent does the instructor's estimate influence the instructor's mark? Does the instructor's estimate of the student's ability relate more strongly to the instructor's mark than to the final examination score?

Consider the following information:

Table 6 - 9

Correlations of Instructors' Ratings of Low-Ability Students' Ability to Meet the Requirements of the Course with Instructors' Marks and Final Examination Grades in Basic 111 and Basic 121.

Type of Grade	Basic 111 Ratings	Basic 121 Ratings
Instructors' Marks	.68	.62
Final Examination Grades	.20	.42

Again there is a greater relationship to the instructors' marks than to the final examination grades.

The correlations shown in this table are higher than corresponding correlations for attitude and effort as indicated in Table 6 - 8 and in the discussion of effort ratings, indicating greater relatedness between this variable and marks than found for the previously discussed variables.

It is entirely possible that the student's performance, his evidence of mastery of the skills and materials, enters into the instructor's appraisal of that student's

lated questions, such as: To what extent does the instructor's estimate influence the instructor's mark? Does the instructor's estimate of the student's ability relate more strongly to the instructor's mark than to the final examination score?

Consider the following information:

Table 6 - 9

Correlations of Instructors' Ratings of Low-Ability Students' Ability to Meet the Requirements of the Course with Instructors' Marks and Final Examination Grades in Basic 111 and Basic 121.

Type of Grade	Basic 111 Ratings	Basic 121 Ratings
Instructors' Marks	.68	.62
Final Examination Grades	.20	.42

Again there is a greater relationship to the instructors' marks than to the final examination grades.

The correlations shown in this table are higher than corresponding correlations for attitude and effort as indicated in Table 6 - 8 and in the discussion of effort ratings, indicating greater relatedness between this variable and marks than found for the previously discussed variables.

It is entirely possible that the student's performance, his evidence of mastery of the skills and materials, enters into the instructor's appraisal of that student's

ability to master those skills and materials. If this is so, there should be a high correlation between the ratings and the final examination grades.

Those correlations were higher than corresponding correlations for the other two variables previously considered and higher than the correlations of A. C. E. Psychological Examination with final examination grades (correlation of .06 for Basic 111 and .09 for Basic 121). However, they are not as high as the correlations between instructors' ratings of ability and instructors' marks.

Only a portion of the rating can be justified as being proven ability as shown by final performance. This indicates that, to some extent, unfulfilled promise of mastery has influenced the instructors' marks. To the extent of this influence, the instructors' marks do <u>not</u> "meaningfully" supplement the final examination grades.

In making their ratings of the students' ability, as was the case in rating attitude, each instructor agreed strongly with his own earlier ratings of the same factor. The correlations between first and second ratings were .85 for Basic 111 and .83 for Basic 121. This indicates that the instructors' judgments were largely determined by mid-term. This, coupled with the evidence of only partial

justification for their ability ratings, strongly suggests that the instructors were allowing early performance by the students to unduly influence the determination of instructors' marks for the quarter's work. If this is true, it has no educational justification.

It is recognized that in the field of human relations, first impressions are important factors in later decisions or actions, and certainly the teacher-pupil relationship is well within that field. However, it is also recognized that first impressions can be, andfrequently are, wrong.

This is not to imply that the instructors' marks are being determined by first impressions, but the evidence strongly suggests that impressions made during the first half of the quarter unduly influence the instructors' marks.

The statements above are based only on findings relative to the low-ability group. It is entirely possible that a similar comparison might lead to different conclusions for a high-ability group. To discover whether this was so, a similar analysis was made of the data pertaining to the "high-ability group" previously described. The findings were as follows:

Table 6 - 10

Correlations of Instructors' Ratings of High-Ability Students' Ability to Meet the Requirements of the Course with Instructors' Marks and Final Examination Grades in Basic 111 and Basic 121.

Type of Grade	Basic 111 Ratings	Basic 121* Ratings
Instructors' Marks	.52	.72
Final Examination Grades	.47	.14

*Basic 121 Correlations are based upon only 24 students.

Since the correlations with Basic 121 ratings are based upon so few cases, the findings might be less representative of the larger group than is the case for Basic 111 ratings. If the difference is actually as great as indicated, teachers in Basic 121 appear to be greatly overestimating the ability of their high-ability students.

The Basic 111 correlations indicate that for the high-ability group, the instructors' judgments of students' ability are in agreement with the evidence of ability shown on the final examination. When this is considered with the findings for the low-ability students it appears that the higher-ability students exhibit a certain performance level and tend to live up to it throughout the quarter, but the low-ability students do not tend to live up to their early performance in the Communication Skills course. However, the instructors, apparently having formed an estimate of ability based upon the student's early performance, continue to hold to that opinion and seemingly allow it to influence their marks.

Again there was very little agreement between ratings by the Basic 111 instructors as a group and the Basic 121 instructors. For the low-ability group, the correlation was .20, and for the high-ability group it was .30. Apparently the early performance in Basic 111 is essentially different from early performance in Basic 121, regardless of the student's ability level. This might be accounted for by the common acceptance that men are more interested in science and do better in it than do women, while women are more interested in communication subjects and do better in them than do the men. Such an explanation would be a convenient device, but the data of this study do not prove it true, although they tend to support it.

Table 6 - 11

The Means and Standard Deviations for Ratings Indicating the Interest in Basic 111 and in Basic 121 Stated by Low-Ability Men and Low-Ability Women.

Subject	Mens' Ratings		Womens' Ratings	
	Mean	S.D.	Mean	S.D.
Communication Skills	4.18	1.00	4.41	1.00
Natural Science	3.67	1.26	3.59	1.38

The ratings were made on a six-point scale. Statistical analysis shows no significant (5% level) difference between the mean ratings by men and women for either subject. However, the trend is in accordance with common belief that women are better in communications work and

men are better in science.

Table 6 - 12

The Means and Standard Deviations for Course Grades Given in Basic 111 and Basic 121 to Low-Ability Men and Low-Ability Women.

Subject	Mens' Grades		Womens' Grades	
	Mean	S.D.	Mean	S.D.
Communication Skills	13.0	3.11	14.8	3.38
Natural Science	13.4	4.42	12.0	4.00

The womens' mean grade was significantly higher than the mens' mean grade in Communication Skills while the mens' mean grade was higher than the womens' mean grade in Natural Science. The means were not significantly different (5% level), however.

The findings of this comparison by itself do not permit a definite statement that the different interest patterns and skills of the two sexes is responsible for the low correlations between the Basic 111 instructors' ratings of students' ability and the Basic 121 instructors' ratings of the same students' ability. However, any difference between the sexes with regard to the skills and materials of the respective courses would tend to reduce the correlation of ability ratings by the two groups of instructors.

In a further study to see whether sex differences might appear in the ratings given, the following findings appeared:

Table 6 - 13

Means and Standard Deviations for Ratings and Scores Received by Low-Ability Men and Low-Ability Women.

Type of Rating or Score	Mens' Ratings or Scores		Womens' Ratings or Scores	
	Mean	S.D.	Mean	S.D.
A.C.E. Test-Retest Average	6.94	2.26	7.44	1.53
111 Instructors' Rating of Effort	6.30	1.87	6.25	1.67
121 Instructors' Rating of Effort	5.04	1.61	5.33	1.79
111 Instructors' Rating of Attitude	7.22	1.79	7.46	1.70
121 Instructors' Rating of Attitude	6.31	1.38	6.33	1.30
111 Instructors' Rating of Ability	5.54	1.64	6.22	1.52
121 Instructors' Rating of Ability	5.47	1.38	5.27	1.67

In only one set of ratings was there a statistically significant (5% level) difference of means. That set of ratings was the Basic 111 instructors' ratings of the students' ability. Apparently the instructors felt that the women students in the low-ability group were more capable of meeting the course requirements. None of the other ratings show such a marked difference. There was no significant difference in mean score for the A.C.E. Psychological Examination test-retest average between the men and women low-ability students.

Jackson[2] reports "The means of the men and women on the Psychological Test are not significantly different." His study was based upon the A.C.E. Psychological Examination Scores received during fall quarter, 1952 by 1687 men and 1296 women, some of whom were students involved in this study.

He also reports that "M.S.C. freshmen women have a significantly higher mean on the English test and the Reading test."[2]

It is entirely possible that the women in this study were superior to the men with regard to meeting the requirements of Basic 111. As an additional bit of evidence, consider the instructors' and final examination grades given in the two subject matter areas:

Table 6 - 14

The Means and Standard Deviations for Scores Received by Low-Ability Men and Low-Ability Women in Basic 111 and Basic 121.

Type of Score	Mens' Scores		Womens' Scores	
	Mean	S.D.	Mean	S.D.
111 Instructors' Score	3.06	2.24	3.14	1.76
111 Final Examination Score	5.02	2.12	5.72	2.43
121 Instructors' Score	6.65	2.59	6.18	2.32
121 Final Examination Score	6.72	2.31	5.86	2.17

[2] Jackson, Robert A., A Report on the Relationship of Orientation Test Scores and First-Term Grade Point Averages, Mimeographed Report, M.S.C. Board of Examiners, 10 Numb. leaves.

Only in the first set of scores is there a significant difference between the means. The instructors in Basic III gave the low-ability women students significantly higher scores than they gave the low-ability men students. This was consistent with their ratings of the students' ability. However, the final examination scores in Basic III failed to show any significant difference between the mean scores for the two groups. The table does show the women receiving a larger mean score, but not significantly so, on their final examinations.

When the Basic III instructors' score was added to the final examination score, the composite, called the course grade, also showed a significant difference between means for the men - women groups (see Table 6 - 10).

In analyzing all of the ratings and grades given by instructors or obtained on final examinations, by comparing those given men to those given women, the following appears evident: In every instance but one, the tendency was for women to rate better than men in Communication Skills and vice versa in Natural Science. However, with but two major exceptions the difference between mean ratings or scores was not statistically significant (5% level). The two exceptions were in the Basic III instructors' ratings of the students' ability and in the Basic III instructors'

marks. Since the final examination scores in Basic 111 failed to show a significant difference between the mens' mean score and the womens' mean score, it is questionable whether such a big difference existed as was indicated by the instructors' marks. If such a difference did not actually exist, then there was a factor entering into the instructors' marks that could not be justified educationally. If it did actually exist, the final examination was too weak to show it.

As a further test of the validity of the instructors' ratings of the students' ability, the ratings given the high-ability group were compared with the ratings given the low-ability group. The findings were as follows:

Table 6 - 15

The Means and Standard Deviations for Ratings of Students' Ability on 114 Low and 91 High-Ability Students by the Basic 111 and Basic 121 Instructors.

	Ratings on Low-Ability Students		Ratings on High-Ability Students	
Type of Rating	Mean	S.D.	Mean	S.D.
111 Instructors' Rating	5.87	1.64	7.86	1.52
121 Instructors' Rating	5.38	1.58	7.50	1.67

In the first case there was no significant difference (5% level) between the mean ratings given the two groups.

The second set of ratings, however, showed a significant difference between the means for the low-ability and high-ability groups. Apparently the Basic 111 instructors could not see sufficient difference between the high- and low-ability students to make that difference significant. The Basic 121 instructors did see significant ability differences between the two-ability groups. In both analyses, however, the t ratio was very close to the critical value.

To see whether the marks given also reflect the same ability judgments, the following information was obtained:

Table 6 - 16

Means and Standard Deviations of Instructors' Scores Given to 117 Low-Ability Students and 23 High-Ability Students by Their Basic 111 and Basic 121 Instructors.

	Scores Given Low-Ability Students		Scores Given High-Ability Students	
Subject	Mean	S.D.	Mean	S.D.
Basic 111	8.59	2.10	10.65	2.26
Basic 121	6.45	2.47	9.92	2.21

In both cases there was a significant difference between the means for the two groups. In giving end-of-quarter marks, both sets of instructors saw differences between the two ability groups that proved to be statistically significant.

When both tables are considered together, it would seem that both the Basic 111 and Basic 121 instructors found differences in ability between the high-ability students as a group and the low-ability students. However, the Basic 111 instructors, in accordance with their general tendency to rate all students favorably, gave ability ratings to the two groups that barely showed "no significant difference" at the 5% level of confidence.

One additional factor was thought to have a contribution to make to the over-all picture of grades and their composition. One of the major problems in teaching is motivation and the student's interest in the subject is considered to be a motivational factor. However, previous studies as reported by Cole[3] failed to show much relationship between the interest that was indicated by the student and the grades that he received. The findings of this study were in agreement with the other studies mentioned. The findings of this study are as follows:

---

[3]Cole, Luella, The Background for College Teaching, Farrar and Rhinehart, Inc., 1940, pp. 226-228.

Table 6 - 17

Correlations Between End of Quarter Scores and Low-Ability Students' Stated Interest in Basic 111 and Basic 121.

Type of Score	Students' Stated Interest in the Course
Basic 111 Instructors' Score	.35
Basic 111 Final Examination Score	-.01
Basic 121 Instructors' Score	.24
Basic 121 Final Examination Score	.25

It would appear that the instructors' score in Basic 111 is most strongly related to the students' stated interest. Possibly this is influenced by the instructors' general tendency to grade favorably and the tendency of the students to sound more interested than they really are.

At most, there is little relationship between the students' stated interest and the end-of-quarter scores that they get.

Cole[4] concludes "One reason for the low relationship between interest and marks may be that once a student is over the 'threshold' of enthusiasm, he is sufficiently stimulated to pass his work; the grade he receives is, then, primarily a function of his ability and his previous preparation for the course."

---

[4]Loc. cit.

CHAPTER VII

CONCLUSIONS AND RECOMMENDATIONS

This study was primarily concerned with the problem of establishing, if possible, pass-fail points for Basic 111 and Basic 121 on a basis of appraisals made of the effort put forth by low-ability students. The general hypothesis proposed by the Basic College Educational Research Committee was that "among a group of students, all low in ability, effort should be a significant factor differentiating those who pass from those who fail."

The findings of this study lead to the conclusion that it is not possible with the instruments and techniques herein described to establish such pass-fail points. One of the reasons for this conclusion was the inability to predict which students were of truly low ability. As a result, the low-ability group included many capable students - students who were not in danger of failing whether they put forth great effort or not.

Another reason for the conclusion was the inability to obtain a pooled effort rating that truly represented the students' study effort. There seemed to be little agreement among the various observers as to the effort that was being made. Even the ratings by the Basic 111 and Basic 121 instructors were little related, though they had many reasons for being similar. Low relationships between the self-

ratings of the students and the ratings by other observers could be readily explained in light of differences in rating instruments, differences in opportunities for observation, differences in experience backgrounds, and differences in objectivity. Other factors such as a general bias toward the more favorable ratings by the students and by their Basic 111 instructors also decreased the extent to which the pooled effort ratings discriminated between the grade levels.

Individual ratings such as the instructors' ratings of students' effort were little better than the pooled effort ratings for the purpose of establishing a pass-fail point. In fact, there was little relationship between instructors' rating of effort and the marks given by those same instructors, where greater correspondence might have been expected. Apparently the Basic 111 and Basic 121 instructors did not allow their opinions as to the students' effort to influence appreciably the marks that they gave the student at the end of the quarter.

While the evidence shown in chapter six suggests that the proposed hypothesis might be valid, gross effort, the over-all quantity of effort, or the physical manifestation of that quantitative effort seemed to have little relation-

ship to either the instructors' marks or the final examination scores. This suggests that more refined instruments must be developed to measure effective effort before a pass-fail point can be established on a basis of student effort.

The conclusion reached in this study relative to the establishment of a pass-fail point was in agreement with the conclusion reached independently by the Basic College Educational Research Committee. They state: "For the present, at least, there appears to be no solution to the problem of grading by attempting to relate a pass-fail point to the effort made by students. It is clear that even the lowest ability students admitted to the college have opportunity to make satisfactory grades - even B's are not uncommon among this group - but the more obvious and overt aspects of effort or application seem unrelated to achievement as appraised either by instructors or by examinations."[1]

The secondary concern of this study was to investigate relative to three factors other than effort, the extent to which the instructors' marks "meaningfully supplement" the final examination scores. The first such factor considered was the student's general attitude toward the course as

---

[1]Basic College Educational Research Committee, The Assignment of Term Marks in the Basic College, Mimeographed Report, Michigan State College, March 1953, 9 Numb. leaves.

appraised by the instructors.

It was found, with respect to this factor, that the instructors' marks did supplement the final examination scores. To the extent that the instructors' appraisals of the students' general attitude toward the course were based upon educationally worth-while traits, habits, or attitudes, the instructors' marks "meaningfully" supplemented the final examination scores.

It was found that there was a stronger relationship between the instructors' ratings of the students' general attitude and the instructors' marks than between those same ratings and the final examination scores. This was taken as evidence that the instructors' marks were supplementing the final examination scores in the determination of the course grade.

The one possible component of the students' general attitude about which this study obtained information was the students' interest in the course. The students' stated interest in the course had little, if any, influence on the ratings given by the instructors relative to the students' general attitude toward the course. Since other studies have shown little relationship between the students' interest in a course and their ability to achieve in the course, the slight relationship found in this study between students'

stated interest and instructors' ratings of students' general attitude might mean that more educationally worthwhile factors were the bases for the instructors' ratings. If this were true, then the instructors' marks did "meaningfully" supplement the final examination scores.

The second factor considered in the study of "meaningful" supplementation of final examination scores by the instructors' marks was the instructors' judgments as to the students' ability to meet the requirements of the course.

It was decided, upon examining the evidence of the study, that the instructors' marks did supplement, for the low-ability students, the final examination scores relative to the factor of student ability. <u>However, they did not "meaningfully" supplement the final examination scores.</u>

There was a stronger relationship between the instructors' ratings of the low-ability students' ability to meet the requirements of the course and the instructors' marks than between those same ratings and the final examination scores. This was taken as evidence that there was supplementation between the instructors' marks and the final examination scores for the low-ability students.

The evidence was contradictory in the case of high-ability students. There was about the same relationship between the Basic 111 instructors' ratings of the students'

ability and the Basic 111 instructors' scores as between those ratings and the Basic 111 final examination scores. However, there was extensive difference between the corresponding relationships for Basic 121. Since this latter set of relationships was based upon only 24 cases, it was discounted as possibly not representing the true situation for all high-ability students.

It appeared that the Basic 111 instructors were basing their ratings of students' ability upon the evidence of performance obtained early in the quarter. The high-ability students seemed to live up to that estimate of their ability and to perform at about the same level on the final examination. The low-ability students, however, seemed to fall short of their instructors' estimates when they took the final examination. The Basic 111 instructors' marks, however, reflecting the instructors' rating of the students' ability to a greater extent than did the final examination scores, indicated that the instructors' judgment as to the students' ability was unduly influencing the instructors' score. It therefore, seemed that the instructors' score contained an element of unfulfilled promise of achievement. If this were so, then to an extent the instructors' marks, while supplementing, were not "meaningfully" supplementing the final examination scores. This non-meaningful supplementation seemed to be greater in extent

than the "meaningful" supplementation of either student effort or student attitude.

In an attempt to discover why there should be so little agreement between the Basic 111 instructors' ratings of students' ability and Basic 121 instructors' ratings of the same students' ability, a study was made of the various ratings and marks or scores to see whether women and men were being appraised or graded differently. It was found that only the Basic 111 instructors were able to see differences in ability and performance between the sexes that proved to be statistically significant. The women received higher mean ratings and marks from the Basic 111 instructors than did the men. This in spite of the fact that the same instructors could see no differences between the sexes on any of their other ratings that were statistically significant and the fact that the Basic 111 final examination scores showed no statistically significant difference between them.

The evidence did suggest that the low-ability women were doing somewhat better work in Basic 111 than the low-ability men and that freshmen women generally were significantly better in English than freshmen men. However, to the extent that the Basic 111 instructors allowed a general opinion that women were superior to men in the skills and comprehensions relating to Communication Skills to influence their marks given to low-ability students, those marks did <u>not</u> "meaningfully" supplement the final examination scores.

The third factor that was studied in considering "meaningful" supplementation between instructors' marks and final examination scores was the students' stated interest in the course.

The findings were such that no general conclusion could be reached with regard to this factor. The students' stated interest in the course was related to about the same extent to instructors' mark and to final examination score in the case of Basic 121. It seemed that no supplementation occurred in this instance.

However, in the case of Basic 111, there seemed to be a greater relationship between students' stated interest and instructors' marks than between students' stated interest and final examination scores. Thus there appeared to be some supplementation; however, it was questionable as to its meaningfulness since there was so little relationship between that stated interest and the instructors' rating of the students' general attitude. The greater relationship might be accounted for by the tendency on the part of the student to appear interested and the tendency on the part of the Basic 111 instructors to give more favorable marks. At least, the evidence was inconclusive.

RECOMMENDATIONS

If the Basic College or some other educational unit desires to establish a pass-fail point on the basis of student effort, it is the recommendation of this study that they first find or develop a better predictive instrument for the identification of truly low-ability students. Careful consideration should be given to the question: Will a prediction of general academic ability be sufficiently discriminating to be useful for selecting persons low in the abilities required for the course under consideration? It is entirely possible that effort will not discriminate between those who should pass and those who should fail unless all of the students are low in the particular abilities that are required for the particular course.

It is also recommended that they develop or find better instruments for obtaining data on the students' study effort. There are two aspects of effort that should be considered, the quantity and the quality. Most of the ratings obtained for this study dealt with quantity. For instance, the dormitory assistants had no way of knowing how effectively students might be studying -- they could only report on a basis of having seen the students go through the motions of studying. The students themselves could report only that they had studied "some" or "quite a bit" or something else

indicating quantity.

It is possible that a quantity rating might be sufficient to discriminate between the grade levels if truly low-ability students are involved. However, instructors' ratings in this study had such a slight relationship to all scores, including the instructors' mark, for a relatively low-ability group that it is questionable whether a sufficiently strong relationship would be found for truly low-ability students.

A quality rating on the students' effort might have a greater chance of discriminating between passing and failing students than a quantitative rating. However, such a qualitative rating will be more difficult to obtain than a quantitative rating of the students' effort.

The Basic College, in its statement of general policy, has stated that the instructors' grades will be determined in the following manner:

> The instructor's grade will be based on such evidence as is, in his judgment, appropriate and is in accordance with policies determined by his department and/or the dean.
>
> Instructors' grades and examination grades should meaningfully supplement rather than duplicate each other.[2]

---

[2]Basic College. Policies and Procedures for Term End Examinations and Term Grades in the Basic College. Unpublished Booklet. Michigan State College. 13 numb. leaves.

The findings of this study indicate that in the case of low-ability students, some instructors' marks appear to be influenced by factors that are not meaningful in light of the general purpose of the Basic College. With this in mind, two recommendations are made.

If the Basic College desires to live up to its statement of general policy, an extensive study should be made to discover what factors are operative in the determination of instructors' marks for all students - not just low-ability students. The factors should then be evaluated as to their meaningfulness in light of the purposes of each department and of the college.

Secondly, an in-service program should be established so that Basic College instructors who are successful in observing and measuring those factors which have been accepted as meaningful might help train others in the necessary techniques.

Should the recommendations of this study be followed, it is likely that the grading system would be improved and that the instructional staff would become more capable of promoting the objectives of general education for which the Basic College was founded.

SELECTED BIBLIOGRAPHY

1. Basic College. Policies and Procedures for Term End Examinations and Term Grades in the Basic College. Unpublished Booklet. Michigan State College. 13 Numb. leaves.

2. Basic College Educational Research Committee. The Assignment of Term Marks in the Basic College. Mimeographed Report. Michigan State College. 13 Numb. leaves. (May, 1952).

3. Basic College Educational Research Committee. The Assignment of Term Marks in the Basic College. Mimeographed Report, Michigan State College, March 1953, 9 Numb. leaves.

4. Board of Examiners, Summary of Basic College Course Grades. Michigan State College, Jan. 21, 1953.

5. Brown, Hugh S. Differential Prediction by the A.C.E. Journal of Educational Research 44:115-121. Oct. 1950.

6. Cole, Luella. The Background for College Teaching. Farrar and Rinehart, 1940, 616 pages.

7. Dressel, Paul L. and Others. Comprehensive Examinations in a Program of General Education. Michigan State College Press, East Lansing. 1949. 165 pp.

8. Froehlich, Gustav J. Academic Prediction at the University of Wisconsin. Journal of American Association of College Registrars 17:65-76. Oct. 1941.

9. Jackson, Robert A. A Report on the Relationship of Orientation Test Scores and First-Term Grade Point Averages, M.S.C. Board of Examiners, 10 Numb. leaves.

10. Leaf, Curtis T. Prediction of College Marks. Journal of Experimental Education 8:303-7. 1940.

11. Osborne, R. Travis, Sanders, W. B. and Greene, J. E. The Differential Prediction of College Marks by A.C.E. Scores. Journal of Educational Research 44:107-115. 1950.

12. Quaid, T. D. D. A Study in the Prediction of College Freshman Marks. Journal of Experimental Education 6:350-75. 1938.

13. Thorndike, E. L. The Nature, Purpose and General Methods of Educational Products. Chapter Two in 17th Yearbook, National Society for the Study of Education, Part 2, p. 16.

14. Thurstone, L. L. and Chave, E. J. The Measurement of Attitude. University of Chicago Press, Chicago, Ill. 1929.

15. Wagner, Mazie E. and Strabel, E. Predicting Performance in College English. Journal of Educational Research 30: 694-9. 1937.

16. Wagner, Mazie E. and Strable, E. Predicting Success in College Physical Sciences. Science Education 19: 4-9, 1935.

APPENDIX

SPECIAL INFORMATION BLANK

nstructor: Department: Course No.

Your Basic College Research Committee, at the request of Dean Erickson, is now ngaged in a study of student achievement. It is of utmost importance to this udy that you make as accurate an appraisal of _____ as you can this time. Please return this rating when completed to the Head of your Departme

Please indicate how this student impresses you by placing a heck (✓) mark at the proper place along the scale line; and lso indicate the basis for your rating.

Check (✓) thos boxes below where you have no basis for a judgment:

tudent's **Ability** to meet the requirements of this course:

Very Low — Average — Very High

Rating is based upon: (Check ✓ One) ( ) Very limited observation ( ) Sufficient observation ( ) Extensive observation

tudent's Effort toward mastering the skills and materials of this course:

Not Trying — Doing a Fair Amount — Working Very Hard

Rating is based upon: (Check ✓ One) ( ) Very limited observation ( ) Sufficient observation ( ) Extensive observation

tudent's general Attitude toward this course:

Very Poor — Ordinary — Very Good

Rating is based upon: (Check ✓ One) ( ) Very limited observation ( ) Sufficient observation ( ) Extensive observation

ow often has this student completed assigned work by the date that it was due?

Never — Occasionally — Frequently — Always

Rating is based upon: (Check ✓ One) ( ) Very limited observation ( ) Sufficient observation ( ) Extensive observation

ow often has this student done more than just the work that was assigned?

Never — Occasionally — Frequently — Always

Rating is based upon: (Check ✓ One) ( ) Very limited observation ( ) Sufficient observation ( ) Extensive observation

he amount of extra help requested by this student:

No Extra Help — Some — Much — An Excessive Amount

f a letter grade were to be placed upon this student's work so far in the course, hat grade would it be? _____

Comments: _____

Use the back of this blank for additional comments if desired.

ar Student:

The Basic College Research committee is making a study to see if the interest at students take in their Basic College courses is related to the studying that ey are doing for those courses. To get an over-all picture, it is necessary to estion a few students that are representative of the general student body. You ppen to be one of the students selected to represent your group. An accurate sponse to the questions below will be greatly appreciated by the Research Committee

The information that you give will be kept strictly confidential and will in no y influence the grades that you receive. Accompanying this sheet is a self dressed envelope in which to seal your reply. Give the sealed letter to your structor before you leave the classroom. He will place it in the Campus Mail thout opening it.

- -

_____ Basic 111 Section_____

ease check (✓) the phrase that most nearly indicates your general interest in the urse named:

Communication Skills Basic 111	Natural Science Basic 121
( ) Not interested at all.	( ) Not interested at all.
( ) Hardly interested.	( ) Hardly interested.
( ) Somewhat interested.	( ) Somewhat interested.
( ) Quite interested.	( ) Quite interested.
( ) Very interested.	( ) Very interested.
( ) Exceedingly interested.	( ) Exceedingly interested.

eck (✓) the words in the sentences below that best describe your studying for your sic courses so far this quarter:

studied ( ) a little / ( ) some / ( ) quite a bit / ( ) very hard earlier in the quarter. I am studying ( ) less / ( ) the same / ( ) more now.

v much time on the average have you spent each week in studying for Basic 111 ____ Basic 121 ____

v much help with your Basic studies have you obtained from other students so far is quarter? Check (✓) the phrase that most nearly indicates the amount.

Basic 111: ( ) None ( ) Some ( ) Quite a bit ( ) A great deal
Basic 121: ( ) None ( ) Some ( ) Quite a bit ( ) A great deal

v much extra help with your Basic studies have you obtained so far this quarter from structors, tutors, dormitory assistants, etc.? Check (✓) the phrase that most arly indicates the amount.

Basic 111: ( ) None ( ) Some ( ) Quite a bit ( ) A great deal
Basic 121: ( ) None ( ) Some ( ) Quite a bit ( ) A great deal

STUDY ATTRIBUTES CHECK LIST

ormitory: Room Number: Assistant:

The Basic College Research Committee is making a study of how student grades are elated to attitudes and study habits. It is of utmost importance to the study that ou make as accurate an appraisal of _____ as can be done at th ime. Information reported on this sheet will be held in strict confidence by the esearch Committee. Please return the completed sheet to your Residence Advisor.

Please check (✔) those expressions which best answer the question as far as the ove named student is concerned. If none of the expressions classify the student t our satisfaction, please write your appraisal in the blanks provided.

o what extent has this student been studying?
Not at all ( ) Very little ( ) _____
About as much as most students do ( )
Quite a bit ( ) Very hard ( ) _____
I have no basis for a judgment at this time. ( )

the amount of studying done by this student has shown any change so far this term hat has been the general nature of that change?
His studying has increased during this term ( ) _____
His studying has decreased during this term ( )
His studying has been sporadic so far this term ( ) _____
I have no basis for a judgment at this time. ( )

hich of the following attitudes does this student have towards any or all of his udies:

Hates { all of his course work. ( ) / some of his course work. ( ) } _____
Dislikes { all of his course work. ( ) / some of his course work. ( ) } _____
Apathetic towards { all of his course work. ( ) / some of his course work. ( ) } _____
Likes { all of his course work. ( ) / some of his course work. ( ) } _____
Loves { all of his course work. ( ) / some of his course work. ( ) } _____
I have no basis for a judgment at this time. ( )

at is your estimate of the number of hours per week that this student spends udying? ___
I have no basis for a judgment at this time. ( )

w much extra help with his studies does this student seek from other students?
None ( ) Some ( ) Quite a bit ( ) A great deal ( )
I have no basis for a judgment at this time. ( )

w much extra help with his studies does this student seek from instructors, tutors, rmitory assistants, etc.?
None ( ) Some ( ) Quite a bit ( ) A great deal ( )
I have no basis for a judgment at this time. ( )

e judgments above are based upon what information or observations?

_____

_____

_____

_____